The Alberta Journal of Educational Research

Vol. II *1956*

THE COMMITTEE ON EDUCATIONAL RESEARCH
Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

THE COMMITTEE ON EDUCATIONAL RESEARCH

Faculty of Education
University of Alberta

VOL. II, No. 3 SEPTEMBER, 1956

CONTENTS

PAGE

EDITORIAL 133

A Qualitative Analysis of Reading Achievement in Edmonton Schools 135
Campbell Young

A Study to Determine the Effect of Transiency on Grade Nine Departmental Examinations Marks 151
Verner R. Nyberg

Financial Assistance Available for Graduate Study in the Field of Education at Canadian Universities 156
Stanley G. Deane

To Teach, or Not to Teach 173
J. D. Aikenhead

An Evaluation of the Effectiveness of a Standardized Test of Language 186
C. H. Sangster

Editorial

RESEARCH AND PHILOSOPHIC INQUIRY

Research is clearly limited by a number of factors: availability of subjects and data, measurability, control. As a result there has been some tendency to distinguish it rigidly from philosophic inquiry. There has been a further tendency to affirm that research is less important than philosophy because it can answer only questions of fact, not of value.

These are unfortunate tendencies. While the central distinction between research and philosophy may be valid enough, it by no means follows that the two must proceed in complete isolation. Research is normally undertaken for the express purpose of providing answers to questions of value, or at least of feasibility. Even if it is not so undertaken, the results of research frequently provoke and assist with answers to questions of value. To require that the objective investigator close his mind to matters of value, or that the deductive thinker dissociate himself from problems of research design and planning would be to push aside needlessly the possibility of broader and better articulated understandings. Surely it is arrogant for the philosopher to say: find me the facts, and I'll tell you what they mean. And just as surely there has been a disposition on the part of some philosophers to adjust the facts to some preconceived scheme of values.

One might almost hope to defend the proposition that those thinkers who have made the greatest contribution have indeed been those whose interests and competence have bridged the two fields.

There is, of course, an historical explanation for the claim that objective research is less important than philosophic inquiry. All science was at one time regarded merely as a branch of philosophy. The rest of the record would seem to show that science began to make its true contribution only when it was emancipated from a position of subservience. And too many of the questions that philosophers still try to answer appear to be answerable, if at all, by objective study and research rather than by sheer ratiocination.

The view here proposed is that of research as an equal partner in the disciplined study of educational problems.

A QUALITATIVE ANALYSIS OF READING ACHIEVEMENT IN EDMONTON SCHOOLS

CAMPBELL YOUNG
Principal, Virginia Park School
Edmonton, Alberta

The Problem

Since reading is probably the most important skill to engage the attention of the school beginner, the progress of pupils in this subject must be under constant examination. Within a few weeks after instruction begins, individual differences in reading become apparent. As the children progress through the grades, differences tend to increase to a startling degree. Wrightstone[1] found a spread of at least five grades in reading ability of 6,581 grade III pupils in New York City. In a similar study of 11,178 grade VIIA pupils, he[2] found an even greater dispersion in reading grade scores, ranging from below grade V to better than grade XII.

Testing of achievement to permit early detection of superiority or weakness in reading is extremely important. Periodically the teacher may make an informal evaluation to check the progress of the class and its individual members. For a more precise measure standardized survey tests and diagnostic tests will serve the purpose better. With this information the teacher may plan more effectively to provide for the needs of both class and individual.

The present investigation is related to Carmichael's[3] study of reading achievement in Alberta. The primary aim of the present study is the analysis of the strengths and weaknesses in reading of 192 grade IV students, and 196 grade VII students, randomly chosen from the Edmonton school system. Attention will be focused on the scores made in reading vocabulary, reading comprehension, their seven component parts, and, to a certain extent, the sub-component elements. The testing instruments were the California Reading Test, Elementary Form AA, and the California Reading Test, Intermediate Form AA. A secondary aim of the investigation is an assessment of the value of the tests.

Experimental Design and Testing Instruments

A survey of reading and language achievement in Alberta schools was conducted in May, 1953. One thousand pupils from each of

---

1J. W. Wrightstone, **Educational Psychology**, p. 485.

2Ibid., p. 486.

3A. Carmichael, **A Survery of the Achievement of Alberta School Children in Reading**, p. 1.

grades IV and VII were included. Two hundred pupils from each grade were randomly selected from each of the five organizational units: large urban, small urban, town, graded rural, and ungraded rural.

Edmonton was selected to provide the large urban sample. Sufficient classes of grade IV and grade VII pupils to fill the required sample were drawn according to a table of random numbers.[4] All pupils in each class drawn were tested. Surplus answer sheets were withdrawn randomly before the tests were scored. With these precautions it was hoped to satisfy the criterion of randomness.

This study will concern itself with the large urban sample drawn from the Edmonton school system.

In the California test, reading vocabulary and reading comprehension are the main fields tested. These are divided and subdivided into component and sub-component parts. The results obtained for each sub-component of the tests were taken from the summary of scores and the Diagnostic Analysis of Learning Difficulties on the last page of the California Reading Test booklet for each grade IV and grade VII student. This information was utilized to study the two main problems set forth for this investigation: reading strengths and weaknesses of Edmonton children and the value of the California Reading Tests.

The methods described will refer to the grade VII sample. However, the organization and analysis of the grade IV results followed, in essence, the same plan.

Mean achievement of the grade VII sample was calculated for the total test, the two main sections, and the components. These mean scores were compared with the California Reading Test norms for grade placement 7.9. By means of a straight line drawn vertically on the diagnostic profile at grade placement 7.9, the norms for the seven components were estimated and compared with the sample's mean component scores. Further analysis of the sample's scores for the main sections and components of the test provided additional data, such as number of perfect scores, range of scores, and median scores.

The percentages of correct responses attained by the sample were calculated in the fifteen components and sub-components of the Diagnostic Analysis of Learning Difficulties. A naive analysis of the results was made.

A study was made of the goodness of the tests in the light of the data secured from the Edmonton samples. The distributions of

[4]E. F. Lindquist, **Statistical Analysis in Educational Research**, p. 119-120.

scores in Vocabulary and Comprehension and the seven components of the diagnostic profile were tested for skewness.

Bloom has named the essential reading skills.

> These skills are speed of reading; knowledge of word meaning, or vocabulary skill; reading for total meaning; reading for central thought; reading for comprehension of special factual information; the ability to read and comprehend and to organize the material that is read; and the ability to read and comprehend and to follow precise directions.[5]

The California Reading Tests would seem to measure reasonably adequately these basic skills.

While some criticism is offered by reading test experts, most of them would seem to agree with Flanagan that teachers will find the California Reading Test "a valuable tool in appraising the progress of pupils with respect to these important skills of vocabulary and reading comprehension."[6]

In her evaluation of the California Test Battery Anastasi makes the following criticism of the Diagnostic Analysis of Reading Difficulties.

> Although of possible help in suggesting specific weaknesses in the individual's mastery of a skill, such an analysis may be quite misleading because of the small number of items involved. The reliabilities of the five major tests at each level appear to be adequate for survey purposes. No reliabilities are reported, however, for the subtests, some of which are very short. As for the functional elements, it is apparent that chance errors of measurement may play a considerable part in the results obtained in individual cases.[7]

The opinions of various reading experts would lead us to conclude that the California Reading Test does measure the fundamental reading skills adequately and reliably. The diagnostic features should prove of use in assaying class achievement. Individual subscores must be regarded as only suggestive.

Related Studies

Competence in the complex skill of reading is so important that teachers are anxious to promote its steady growth in their students. Russell[8] believes that evaluation is one of the most effective ways of preventing crystallization or deterioration of an instructional programme.

Tiegs[9] compared the reading scores of approximately 230,000 elementary pupils in seven states of the Stanford and Progressive

---

5M. E. Bloom et al., **Effective Reading Instruction**, p. 246.

6J. C. Flanagan, **Fourth Mental Measurement Yearbook**, p. 568.

7A. Anastasi, **Psychological Testing**, p. 469.

8D. H. Russell, "Evaluation of Pupil Growth In and Through Reading", **N.S.S.E. Forty-Eighth Yearbook**, Part 11, 1949, p. 285.

9E. W. Tiegs, **Progressive Education** (January, 1952), p. 106.

Achievement Tests before and after 1945. The scores indicated a slight gain in the achievement-level of pupils despite the fact that more children who formerly failed and dropped out were remaining in school, and the average chronological age was decreasing.

In a recent study, Coull[10] found that Alberta children in grades IV and VII were respectively superior and equal in reading achievement to the norms of the California Reading Tests, Form AA, in terms of grade scores. He also found an indication of deceleration in growth of reading ability between grades IV and VII.

Strengths and Weaknesses in The Reading of Edmonton Children

Grade IV data

One of the main purposes of this study was to secure accurate, objective and unbiased information about the reading performance of the Edmonton sample as measured by the California Reading Tests. Tables I to IV will present findings with respect to grade IV. Tables V to VIII will provide similar information about grade VIII.

Table I compared mean achievement of the grade IV sample with the norms of the California Test for Vocabulary, Comprehension, and Total Reading. It is to be noted that the Edmonton data were secured in the ninth month in grade IV. As a result, achievement will be compared with the test norms provided for grade 4.9.[11]

TABLE I

COMPARISON OF MEAN RAW SCORES FOR GRADE FOUR IN VOCABULARY, COMPREHENSION, AND TOTAL READING WITH CALIFORNIA READING TEST NORMS FOR GRADE PLACEMENT 4.9

Reading Skills	Mean of Edmonton Raw Scores	California Raw Score Norms	Difference
Vocabulary	70.48	66	4.48
Comprehension	26.25	21	5.25
Total Reading	96.73	87	9.73

The comparison in Table I indicates that the grade IV sample exceeded the norms in Vocabulary, Comprehension, and Total Reading.

10W. H. Coull, **A Normative Survey of Reading Achievement of Alberta Children in Relation to Intelligence, Sex, Bilingualism and Grade Placement**, p. 57.

11Ernest W. Tiegs and Willis W. Clark, **Manual, California Reading Test—Elementary**, p. 20.

Table II draws a comparison of the means of the raw scores obtained by the grade IV sample in the seven components of the California Reading Test with the estimated norms for these components. The estimates were secured by drawing a vertical line through grade placement 4.9 on the diagnostic profile on the last page of the test booklet.

TABLE II

COMPARISON OF MEAN RAW SCORES IN THE SEVEN COMPONENTS FOR GRADE FOUR WITH THE NORMS ESTIMATED FROM THE DIAGNOSTIC PROFILE AT GRADE PLACEMENT 4.9

Reading Skills	Means of Edmonton Raw Scores	Estimated California Norms	Differences
Vocabulary			
Word Form	23.66	23.20	.46
Word Recognition	18.96	17.60	1.36
Opposites	13.99	12.60	1.39
Similarities	13.91	11.50	2.41
Comprehension			
Following Directions	7.18	5.90	1.28
Reference Skills	6.46	5.50	.96
Interpretations	12.61	9.00	3.61

The mean scores of the Edmonton sample exceeded the estimated California norms in all seven components; the superiority in Interpretations being particularly striking. From this greatest display of superiority the difference of mean raw scores ranged downward to .46 in Word Form, the smallest difference noted.

Table III presents more data concerning scores obtained by the grade IV sample in the two major sections and the seven components of the test.

A wide range in ability in the reading skills is clearly indicated. Reading grade placement levels in the seven components, estimated from the diagnostic profile, vary from below 2.0 to above 9.0. The median scores are in every case higher than the mean scores obtained by the group for the components shown in Table II. The fact that the groups' mean and median scores exceeded the California component norms seems to indicate that the sample found the test relatively easy.

Perfect scores in three of the component were comparatively numerous. Of the 192 pupils in the sample, perfect scores were made by 76 pupils in Word Form, by 106 in Word Recognition, and

by 19 in Following Directions. It would appear that the tests are inadequate in that they fail to differentiate the ability levels of the many pupils attaining perfect scores.

TABLE III

SCORES OBTAINED BY THE GRADE FOUR SAMPLE FOR THE MAJOR SECTIONS AND THE SEVEN COMPONENTS OF THE DIAGNOSTIC PROFILE

Reading Skills	Possible Score	Number of Perfect Scores N = 192	Range of Scores	Median
Major Sections				
Vocabulary	90	0	33-88	71.00
Comprehension	40	0	10-39	26.17
Components				
Word Form	25	76	11-25	24.20
Word Recognition	20	106	1-20	19.59
Opposites	23	0	0-20	14.26
Similarities	22	1	3-22	14.10
Following Directions	10	19	0-10	7.36
Reference Skills	10	6	1-10	6.55
Interpretations	20	1	3-20	12.62

Table IV shows the percentages of correct responses attained by the grade IV sample from the total possible scores obtainable by the group in the 19 components and sub-components of the Diagnostic Analysis of Learning Difficulties. Ratings are given for 192 pupils.

The percentage of correct responses scored by the sample in the components and sub-components of the diagnostic analysis show, in a general way, the elements of the reading skills which are easy for grade IV pupils. However, these scores are merely suggestive for there are only a few test items in each sub-component of the diagnostic analysis. The relatively low percentages of correct responses obtained by the sample in the sub-components, Definitions and Directions, Use of Index, Organization of Topics, and Sequence of Events, might point to weakness in these skills. To improve scores in reading comprehension teachers might be advised to provide more training in these areas. It should be remembered, however, that the position of the two sub-components at the end of the test might account, in part, for the relatively poor scores obtained.

Grade VII data

Table V presents a comparison in Vocabulary, Comprehension, and Total Reading of the means of the raw scores of the grade VII

sample with the California Test norms provided for grade 7.9.[12] It is again to be noted that the Edmonton data were secured in the ninth month of grade VII.

TABLE IV

PERCENTAGE RATINGS OF CORRECT RESPONSES FOR GRADE FOUR STUDENTS FOR THE COMPONENTS AND SUB-COMPONENTS OF THE DIAGNOSTIC ANALYSIS OF LEARNING DIFFICULTIES

Reading Skills		Percentages of Correct Responses
Components	Sub-Components	N = 192
Word Form	Lower Case Words	96.74
	Capitals	95.31
	Miscellaneous Type Faces	88.80
Word Recognition	Gross Differences	98.44
	Initial Sounds or Endings	94.12
Opposites	Basic Vocabulary	60.82
Similarities	Basic Vocabulary	63.26
Following Directions	Simple Directions	78.39
	Direct Simple Choice	77.26
	Definitions and Directions	48.70
Reference Skills	Parts of Book	77.08
	Alphabetizing	59.64
	Table of Contents	72.40
	Use of Index	51.74
Interpretations	Topics or Central Idea	85.16
	Directly Stated Facts	81.86
	Making Inferences	69.90
	Organization of Topics	37.67
	Sequence of Events	34.11

TABLE V

COMPARISON OF MEAN RAW SCORES FOR GRADE SEVEN IN VOCABULARY, COMPREHENSION, AND TOTAL READING TEST NORMS FOR GRADE PLACEMENT 7.9

Reading Skills	Mean of Edmonton Raw Scores	California Raw Score Norms	Difference
Vocabulary	58.78	57	1.78
Comprehension	35.68	35	.68
Total Reading	94.45	92	2.45

[12]Ernest W. Tiegs and Willis W. Clark, **Manual, California Reading Test—Intermediate, p. 16.**

Comparison of the achievement of the grade IV and grade VII samples with their respective test norms in Vocabulary, Comprehension, and Total Reading, reveals an apparent decline in the relative superiority of the grade VII pupils.

Table VI makes a comparison of the means of the raw scores obtained by the grade VII sample in the seven components of the California Reading Test with the estimated norms for these components. Following the same method used with grade IV, estimates of normal performance in the ninth month of grade VII were made from the diagnostic profile on the back page of the test booklet.

TABLE VI

COMPARISON OF MEAN RAW SCORES IN THE SEVEN COMPONENTS FOR GRADE SEVEN WITH THE NORMS ESTIMATED FROM THE DIAGNOSTIC PROFILE AT GRADE PLACEMENT 7.9

Reading Skills	Mean of Edmonton Raw Scores	Estimated California Norms	Difference
Vocabulary			
Mathematics	15.80	15.90	— .10
Science	13.53	12.30	1.23
Social Science	15.96	14.80	1.16
General	13.54	13.30	.24
Comprehension			
Following Directions	7.39	6.30	1.09
Reference Skills	9.31	9.20	.11
Interpretations	18.98	17.80	1.18

The Edmonton sample mean scores exceeded the estimated California norms in all components but Mathematics where the sample deficiency was slight. The remaining six components showed relatively small differences in favor of the Edmonton sample, ranging up to 1.23 for Science.

Table VII presents additional data about scores obtained by the grade VII sample in the two major sections and the seven components of the test.

Table VII shows clearly that a wide range of reading ability exists in the grade VII sample. Reading grade placement levels in the seven components, estimated from the diagnostic profile, vary from below 4.0 to above 11.0. The median scores are in every case higher than the mean scores obtained by the group for the seven components. The sample's mean scores were higher than the estimated norms for the test with the previously noted exception,

Mathematics. The slight superiority exhibited by the Edmonton sample, in relation to the estimated norms for the seven components, would seem to indicate that the skills measured were adequately mastered.

TABLE VII

SCORES OBTAINED BY THE GRADE SEVEN SAMPLE FOR THE MAJOR SECTIONS AND THE SEVEN COMPONENTS OF THE DIAGNOSTIC PROFILE

Reading Skills	Possible Score	Number of Perfect Scores N = 196	Range of Scores	Median
Major Sections				
Vocabulary	90	0	15-86	58.92
Comprehension	55	0	11-50	36.02
Components				
Mathematics	22	2	3-22	16.53
Science	23	0	0-21	13.50
Social Science	23	14	0-23	16.39
General	22	1	5-22	13.76
Following Directions	10	34	1-10	7.71
Reference Skills	15	1	0-15	9.50
Interpretations	30	0	7-29	19.00

The sample's mean raw score in Mathematics was 15.80 compared with a estimated norm of 15.90. However, the sample's median score of 16.53 was higher than the norm. This seems to indicate that a majority of the sample earned higher ratings than the norm but a portion of the sample with relatively low scores depressed the mean score below the norm.

In the sample of 196 pupils, perfect scores were made by 14 in the component, Social Science, and by 34 in Following Directions. There seems to be too great a percentage of perfect scores in these components with a resulting failure to differentiate the levels of skill among the more able students. The inability of the test to determine the upper achievement limit of these pupils depressed the mean scores of the group.

Table VIII shows the percentages of correct responses attained by the grade VII sample from the total possible scores obtainable by the group in the 15 components and sub-components of the Diagnostic Analysis of Learning Difficulties. Ratings are given for 196 pupils.

The percentages of correct responses scored by the sample in the components and sub-components of the diagnostic analysis suggest the elements of reading skills which were inadequately mastered.

As with the grade IV results, these findings represent merely possible deficiencies in reading skill. The small number of items in the sub-tests render the reliability of their results open to question. However, the relatively low percentages of correct responses obtained by the sample in the sub-components, Selecting References, Organization of Topics, and Sequence of Events, hint at weakness in these skills. It should be noted that the grade IV sample was weak in Organization of Topics and Sequence of Events. The fact that these two sub-components appear at the end of the test may have some bearing on the low scores obtained in them by the sample.

TABLE VIII

PERCENTAGE RATINGS OF CORRECT RESPONSES FOR GRADE SEVEN STUDENTS FOR THE COMPONENTS AND SUB-COMPONENTS INCLUDED IN THE DIAGNOSTIC ANALYSIS OF LEARNING DIFFICULTIES

Reading Skills		Percentages of Correct Responses
Components	Sub-components	N = 196
Mathematics		71.71
Science		58.81
Social Science		69.41
General		61.53
Following Specific Directions	Simple Choice	76.28
	Definition and Directions	73.28
Reference Skills	Parts of Book	68.78
	Table of Contents	88.27
	Use of Index	83.84
	Selecting References	31.84
Interpretation of Meanings	Topic or Central Idea	73.47
	Directly Stated Facts	71.72
	Inferences	68.80
	Organization of Topics	45.28
	Sequence of Events	38.78

Goodness of the Tests

It seems reasonable to assume that the California Reading Tests are constructed in such a way that, if a large random sample is tested, the frequency distribution of the resulting scores will closely resemble the normal curve. Should there be any significant deviation from normality when a large, randomly chosen sample is tested, it may indicate the existence of defects in the test, undetected selective factors in the sampling, or poor administration of the test and its results.

Carmichael[13] outlined the methods used to obtain a random sample in the provincial survey. Standard procedures were followed. No effort was spared to secure uniformly excellent administration and handling of the tests and their results.

Grade IV sample

Table IX summarizes the results of tests for skewness of the scores for grade IV in Vocabulary and Comprehension and the seven components of the diagnostic profile.

TABLE IX

SKEWNESS OF SCORES OF VOCABULARY AND COMPREHENSION SECTIONS AND THE SEVEN COMPONENTS OF THE DIAGNOSTIC PROFILE OF THE TEST FOR GRADE FOUR

Reading Skills	S.	S.E.	C.R.	Level of Significance
Sections				
Vocabulary	—1.23	.94	—1.31	.....
Comprehension	.22	.54	.41	.....
Components				
Word Form	— .78	.14	—5.57	.01
Word Recognition	— .85	.12	—7.08	.01
Opposites	— .40	.40	—1.0	.....
Similiarities	— .45	.40	—1.13	.....
Following Directions	— .39	.19	—2.05	.05
Reference Skills	— .12	.18	— .67	.....
Interpretations	.30	.33	.91	.....

NOTE: S—Skewness; S.E.—Standard Error of Skewness; C.R.—Critical Ratio.

The measurements of skewness in Vocabulary and Comprehension revealed non-significant departures from normality. However, when the scores for the seven components were examined, evidence of significant skewness in their distributions was apparent in several instances.

The normality of the distritbutions for Vocabulary and Comprehension for the grade IV sample indicates that, on the whole, the Elementary Test, Form AA, meets this requirement.

The distributions of scores for components, Word Form and Word Recognition, were skewed negatively at the .01 per cent level of significance. Asymmetry in these cases may result from the lack of more advanced items in the tests with the result that little dis-

[13]A. Carmichael, op. cit., p. 8.

crimination among superior students would be possible. Exclusion of these items would prevent an extension of the curve at the upper end of the distribution. In a somewhat lesser degree, the same criticism was true for the component, Following Directions, for which the distribution of scores was skewed negatively at the .05 per cent level of significance.

Negative skewness of the distributions of scores for Word Form, Word Recognition, and Following Directions in the Elementary Test, Form AA, was sufficient to jeopardise the value of the test as an effective diagnostic instrument for Edmonton schools.

Except for Interpretations, the components were skewed negatively, but represented no significant deviation from normality.

The distribution of Total Reading scores approached the normal curve. This tendency characterizes most good tests. However, only four of the seven components exhibited normality in their distributions of scores. The divergence from normality on the part of the remaining components indicates two possibilities: the items of these components are weak or are insufficient in number.

Significant skewness of a negative nature indicated those parts of the test that were too easy for the Edmonton sample. This suggests that the curriculum of the pupils on whom the California Reading Tests were standardized does not require the Edmonton schools' level of achievement. Therefore, the tests will deal unfairly with superior members in any group of Edmonton pupils.

Grade VII sample

Table X summarizes the results of tests for skewness and their significance for the grade VII distributions of scores in Vocabulary and Comprehension and the seven components of the diagnostic profile.

The normality of the distributions of scores for the Vocabulary and Comprehension sections attained by the grade VII sample suggests that the Intermediate test, on the whole, is good. The measurement of Skewness in Vocabulary and Comprehension represent departures from normality that are relatively mild and non-significant.

The range of difficulty covered by the test items in the components of Science, Social Science, General, Reference Skills and Interpretations, appeared to suit the sample, for the resulting distributions of scores exhibited no significant deviation from normality.

The other component parts may not be so well constructed, however. The extreme negative skewness of the distributions of scores for the components, Mathematics and Following Directions, suggests

that these tests were far too easy for the sample and failed to discriminate among the abler members of the group. .

TABLE X

SKEWNESS OF SCORES OF VOCABULARY AND COMPREHENSION SECTIONS AND THE SEVEN COMPONENTS OF THE DIAGNOSTIC PROFILE OF THE TEST FOR GRADE SEVEN

Reading Skills	S.	S.E.	C.R.	Level of Significance
Sections				
Vocabulary	— .94	1.19	— .79	.....
Comprehension	— .64	.81	— .79	.....
Components				
Mathematics	—1.23	.29	—4.24	.01
Science	.21	.32	.66	.....
Social Science	— .54	.45	—1.20	.....
General	— .23	.27	.85	.....
Following Directions	— .59	.21	—2.81	.01
Reference Skills	— .28	.23	—1.22	.....
Interpretations	— .11	.50	— .22	.....

NOTE: S—Skewness; S.E.—Standard Error of Skewness; C.R.—Critical Ratio.

The writer agrees with the authors who point out in the test manual that "because of the limited number of items (10 to 30), scores on the seven sections of each test should be used only as guides to indicate the presence of student difficulties in the seven major diagnostic areas.[14] Nevertheless, when measuring the achievement of 196 pepole, it seems reasonable to expect less deviation from normality in the distribution of scores than has been shown in the components, Mathematics and Following Directions,

Conclusions and Recommendations

Almost without exception, the normal class in a city school contains pupils whose reading skills vary all along the scale from good to poor. The periodic testing of pupils' reading strengths and weaknesses is a useful practice, particularly if teachers and administrators are ready and able to provide improved reading instruction based on the findings of the tests. Survey and diagnostic tests are designed to assist in the planning of a helpful reading programme in the classroom.

This investigation has sought to answer two questions related to

[14]Tiegs and Clark, op. cit., p. 4.

the reading ability of two samples of Edmonton school pupils tested by the California Reading Tests. A summary of the conclusions and some recommendations follow.

Strengths and weaknesses in the reading of Edmonton children

(1) The mean scores of both grade IV and grade VII sample groups exceeded test norms for Vocabulary, Comprehension, Total Reading, and the estimated norms for the seven component parts, with the exception of grade VII Mathematics. The margin of superiority was less in grade VII than in grade IV.

(2) Scores made by each sample in the seven components of their respective tests revealed a variation in individual grade placement scores of more than seven grades. The grade IV sample varied from below grade 2.0 to above grade 9.0. The grade VII sample varied in grade placement from below 4.0 to above 11.0. Both samples contained pupils whose upper limits of reading ability were not adequately measured. The true range of ability may have varied more than was indicated by the test results.

(3) The relatively low percentages of correct responses obtained by the grade IV sample in the sub-components, Definitions and Directions, Use of Index, Organization of Topics, and Sequence of Events, might point to weaknesses in these skills. Perhaps grade IV teachers in Edmonton should be advised to provide more training in these areas of reading comprehension.

(4) The grade VII sample achieved relatively low percentages of correct responses in the sub-components, Selecting References, Organization of Topics, and Sequence of Events. This suggests possible weaknesses in these skills and hints that Edmonton grade VII teachers might provide more training to improve their students' techniques in these skills.

(5) It should be noted that both samples were weak in Organization of Topics and Sequence of Events.

Goodness of the tests

(1) The normality of the distributions of scores for the Vocabulary and Comprehension sections attained by the two samples suggests that the tests, on the whole, are good.

(2) Departures from normality in the distribution of scores for certain of the components of the two California tests suggest a cautious approach in the use of the tests as diagnostic instruments. Significant skewness of distributions of scores was noted in the components, Word Form Word Recognition, and Following Directions in the California Reading Test, Elementary, Form AA. In the California Reading Test, Intermediate, Form AA, skewness was

significant in the components, Mathematics vocabulary and Following Directions.

(3) The findings in relation to the Edmonton sample indicate that the value of the tests seems to lie mainly in their use in survey programmes with the diagnostic components and sub-components being helpful in suggesting possible strengths and weaknesses in reading ability at the individual or group level.

Recommendations

(1) The relatively low percentage of correct responses obtained by the grade IV sample in the sub-components, Definitions and Directions, Use of Index, Organization of Topics, and Sequence of Events, might point to weaknesses in these skills. Therefore, grade IV teachers should perhaps be advised to provide more training in these areas in order to improve the reading comprehension of their pupils.

(2) The grade VII sample achieved relatively low percentages of correct responses in the sub-components, Selecting References, Organization of Topics, and Sequence of Events. This suggests possible weakness in these skills and hints that Edmonton grade VII teachers might be well advised to place more emphasis on these specific areas of the reading programme.

(3) An examination of scores obtained in the component tests revealed that in each sample, the range of reading ability extended over more than seven grades. While it is difficult to attend to the needs of individuals in a large class with such divergent levels of ability, teachers and administrators are obligated to plan their reading programme with these individual differences in mind.

(4) In the components and sub-components of the tests where many pupils made perfect scores, the provision of more difficult items to ensure measurement of the upper ability levels of Edmonton children seems required.

(5) Local norms should be developed for the main sections and the seven components of the California Reading Tests.

(6) The California Reading Tests should be used as survey tests and as a means of suggesting the reading strengths and weaknesses of groups and individuals. It has not been proven adequate as a precise diagnostic instrument.

BIBLIOGRAPHY

Anastasi, A. *Psychological Testing.* New York, The Macmillan Company, 1954.

Bloom, M. E. et al. *Effective Reading Instruction* (Second Edition). New York, McGraw-Hill Book Co., Inc., 1951.

Burkhart, K. H. "An Analysis of Reading Abilities" *Journal of Educational Research* XXXVIII (February, 1945), pp. 430-439.

Carmichael, A. *A Survey of the Achievement of Alberta School Children in Reading.* M.Ed. Thesis, University of Alberta, 1954.

Coull, W. H. *A Normative Survey of Reading Achievement of Alberta Children in Relation to Intelligence, Sex, Bilingualism, and Grade Placement.* M.Ed. Thesis, University of Alberta, January, 1956.

Cronbach, Lee J. *Essentials of Psychological Testing.* New York, Harper and Brothers, 1949.

Flanagan, John C. *The Fourth Mental Measurements Yearbook.* New Jersey, The Gryphon Press, 1953.

Garrett, H. E. *Statistics in Psychology and Education.* New York, Longmans, Green and Co., Inc., 1953.

Jordan, A. M. *Measurement in Education.* New York, McGraw-Hill Book Co., Inc., 1953.

Lindquist, E. E. *Statistical Analysis in Educational Research.* Boston, Houghton Mifflin, 1940.

Russell, D. H. "Evaluation of Pupil Growth In and Through Reading." *N.S.S.E. Forty-Eighth Yearbook,* Part II, 1949.

Tiegs, E. W. *Progressive Education* (January, 1952).

Tiegs, E. W. and Clark, W. W. *Manual, California Reading Test, Intermediate.* Los Angeles, Calfifornia Test Bureau, 1950.

Tiegs, E. W. and Clark, W. W. *Manual, California Reading Test, Elementary.* Los Angeles, California Test Bureau, 1950.

Wrightstone, J. W. *Analysis of Intelligence and General Achievement* (Third Grade Classes) and *Analysis of Results Survey of Intelligence and Reading Ability in 7A Classes.* In Leonard Wood, Educational Psychology (Third Edition), New York, Prentice Hall Inc., 1951.

A STUDY TO DETERMINE THE EFFECT OF TRANSIENCY ON GRADE NINE DEPARTMENTAL EXAMINATIONS MARKS

VERNER R. NYBERG
Provost, Alberta

The Problem

The discovery of oil near Leduc, Alberta, in 1947 gave the petroleum industry of this province an impetus that has not yet subsided. Subsequent exploration and discoveries brought in their wake features heretofore unknown. One such feature was the advent of transient workers and their families.

The impact of this influx on the communities concerned was one from which the school did not escape.

As time went on, it became apparent that transiency was not a feature that would soon pass. Each new oil field discovery meant the arrival and departure of the exploration men, then the entry of drilling crews of many companies who were followed by pipeline gangs, service-rig crews, maintenance and supply men and production personnel. In addition, there were those who conducted their own business ventures and followed the oil development to take advantage of the better opportunities for finding work.

The transiency connected directly with oil is just one phase of the problem in Alberta. The development of industry that followed the discovery of petroleum has increased the number of transients, and, judging from appearances, this situation will likely continue for some time.

Census figures indicate that since 1947 there has been a great increase in transiency in Alberta, and it follows that problems arising from the shift in population are greater than before. There seem to be no indications at present that movements of population are going to decrease in the near future.

Purpose of this Study

Transiency could be the cause of many problems pertinent to the school. Any research conducted with limited time and resources must necessarily be carried out within a small and definite range.

The purpose of this study was to determine the effect of transiency on the marks received by grade IX pupils in Alberta on the Departmental Examinations of June, 1953. To state the problem in another way, this study was made to determine whether the

changes resulting from transiency, such as (1) becoming adapted to a new school situation, (2) becoming adapted to a new home life, (3) coming into contact with a new social group, will enhance pupil achievement in grade IX, have a detrimental effect, or have no effect upon it.

The problem is *not* to find out whether the transient pupil, on the average, is a better or poorer student than the average non-transient, but to discover whether he gains or loses or is not affected by a transfer from one school to another.

Procedure

To complete this study, information from a considerable number of grade IX pupils was required. This information included personal data, such as name and age; information concerning schools attended in grades VII to IX and the number of months in each; and the marks obtained by the pupils in each of the subjects included in the grade IX final examinations conducted by the Department of Education.

About 2,500 forms were printed and distributed, chiefly to pupils of schools in Central Alberta. Approximately two thousand cards, containing the personal data and history of schools attended, were returned by the end of June 1953.

The next step was to secure from the records of the Department of Education the marks received by the pupils in the grade IX final examinations. The attendance figures supplied to the Department office by the teachers were also recorded. Since, at this stage, it had not yet been decided exactly which pupils were to be included in the study, it was necessary to copy the marks received by all the pupils who had completed the cards previously mentioned. The personal data supplied by the pupils on the form were checked against the information given to the Department. Cases where records were incomplete were not considered further.

For each pupil so far classified as a transient a corresponding non transient was found. The pairing was done with considerable care because the extent to which the variables of sex, age, mental ability, attendance and type of school attended could be controlled depended upon how closely each transient could be matched with a non-transient. Members of a pair had to be of the same sex, of the same age (within seven months, if possible), of nearly the same achievement in the tests of mental ability (within five percentile points), and of nearly the same attendance (within fifteen days, if possible). In addition, pairs were chosen so that the types of schools attended in each case were nearly the same. The matching, with

respect to this factor, consisted mainly in seeing to it that no pupil who had spent most of his time in a city school was paired with one who had attended mainly a town, village or centralized school. Other factors were set as nearly equal as the choice permitted. For example, pupils from separate schools were paired with each other as often as could be arranged. No set limits with respect to these factors were devised because it was considered unnecessary to do so.

When the pairs, one hundred in all, were established they were closely checked again. To guard against the possibility of giving one group or the other an advantage resulting from small differences being unconsciously but consistently allowed when each of the pairs was constituted, a number of grand totals for each group was struck. The scores on the general test, the attendance in days, the age in months, the time in months spent in each type of school attended in grades VII to IX, and the time in grade IX only, were totalled for each group and differences noted. Changes in the pairs were made so that the grand totals were more nearly the same. Tables showing personal data for each group were set up so that final comparison could be made. Table 1 gives a summary of the statistical analysis of the attendance figures for the transient and non-transient groups and also the age and achievement in the General Test.

TABLE I

RESULTS OF STATISTICAL ANALYSIS OF ATTENDANCE AND AGE FIGURES AND OF PERCENTILE RANKS ATTAINED BY TRANSIENT AND NON-TRANSIENT PUPILS ON THE DEPARTMENTAL GENERAL TEST

Item Analysed	Means		Differences in Means ($M_2 - M_1$)	Standard Deviations		Differences in Deviations	Critical Ratios Pertaining to—	
	M_1 Transient	M_2 Non-Transient		Transient	Non-Transient	Standard	Means	Standard Deviations
Attendance (days)	172.4	174.2	1.80	12.99	12.02	0.97	1.01	0.77
General Test (%ile Ranks	55.28	55.40	0.12	18.05	17.73	0.32	0.05	0.18
Age (years)	15.30	15.30	0.00	8.84	7.84	1.00	0.00	1.43

In view of the low critical ratios found in Table I, it was concluded that there is no significant difference in the average number of days attended by each group, nor is the variability of the attendance of one group significantly different from that of the other. In the case of the General Test and of age, it is readily seen that the groups

are nearly identical. This applies to the variability as well as the mean score. It was concluded that any differences between transient and non-transient pupils in performance on the grade IX examinations is not attributable to differences in attendance and age nor to unequal mental abilities.

Results

Table II gives a summary of the statistical analysis of the marks received by the transient and non-transient pupils on the final examinations. The figures reveal the extent and significance of any differences between the two groups.

In the case of each grade IX subject, a null hypothesis that there was no real difference between the transient pupils and the non-transient pupils was set up. In the light of the critical ratios derived for each subject the null hypothesis was then accepted or rejected.

It should be borne in mind, while studying the figures in Table II, that a critical ratio of 1.98 or over, as applied to the arithmetic means, would indicate a real difference, at the .05 level of con-

TABLE II

RESULTS OF STATISTICAL ANALYSIS OF MARKS RECEIVED BY BOTH TRANSIENT AND NON-TRANSIENT GROUPS IN DEPARTMENTAL EXAMINATIONS

Subject	Means		Differences in Means $M_2 - M_1$	Standard Deviations		Differences in Standard Deviations $CR_1 - CR_2$	Critical Ratios Pertaining to—	
	Transient M_1	Non-Transient M_2		Transient CR_1	Non-Transient CR_2		Means	Standard Deviations
Reading	55.83	55.81	(—) 0.02	16.09	16.07	0.02	0.01	0.01
Literature	28.22	28.18	(—) 0.04	8.46	8.28	0.18	0.03	0.22
Language	53.34	55.33	1.99	17.15	15.59	1.56	0.85	0.95
Social Studies	48.95	53.89	4.94	17.37	15.88	1.49	2.09	0.89
Mathematics	51.40	54.26	3.86	16.59	17.48	(—) 0.89	1.59	0.52
Science	48.81	53.00	4.19	18.69	16.02	2.67	1.70	1.53

fidence, between the transient and non-transient groups. A critical ratio approaching 1.98 in the case of the standard deviations would indicate that there is a great difference in the variability of each group and would cast doubt on any conclusions reached.

The critical ratios given in Table II suggest that only in the case of social studies must the null hypothesis that the two groups are drawn from the same population be rejected. A difference between the means as great as that shown is significant at the .05 level of confidence. In the case of reading, literature, language, mathematics and science, the null hypothesis must be accepted. The difference between the means in each case is attributable to chance or errors of sampling.

In every instance, the difference between the two standard deviations is not significant. It may be assumed, therefore, that the variability of the marks in each subject is much the same.

Findings

In the light of the facts brought out in this study, a number of conclusions may be made. First, the figures in Table II lead one to believe that when a pupil in grade IX transfers from one school to another his final mark in social studies will be lower than if he had not changed schools. One explanation of this finding might be that it is difficult for a teacher who receives a transient pupil to determine how much pertinent information the student absorbed in the previous school. Furthermore, because of the nature of the material, the various sections of the subject are not dealt with in the same sequence in all schools. Some of the transient students, therefore, will be taught the same unit twice but will receive no instruction in another section of the course. Finally, it is more difficult to organize remedial work in social studies than it is in some of the other subjects.

A second conclusion suggested by the data in Table II is that this study did not reveal a significant relationship between transiency and performance in any subject other than social studies. It might be noted at this point that in all subjects containing an organized body of knowledge, the differences, however small, between the transient and non-transient pupils are in favor of the latter. This might indicate a trend which could be verified only by further study.

Finally, it seems reasonable to make the generalization that grade IX pupils in Alberta lose in some fields by moving from one school to another during the term. However, the amount of lowering of the marks, even in the subject where it is greatest, is not enough to be alarming. The better performance by the transients in reading and literature is very small. Suffice it to say that there is a possibility that transiency benefits a pupil in these fields.

FINANCIAL ASSISTANCE AVAILABLE FOR GRADUATE STUDY IN THE FIELD OF EDUCATION AT CANADIAN UNIVERSITIES

Stanley G. Deane
Eastglen Composite High School

The Problem

The report of the First Canadian Conference of Deans of Education and Professors of Educational Administration and Supervision, 1955, indicated three needs emerging from conference discussion.[1] Of these the third was a need for greater financial aid to graduate students in educational administration and supervision.

Preliminary to any action that might be taken there was need for a summary of the kinds and amounts of available aid in the form of scholarships, fellowships, bursaries, loans, assistantships, special grants, and paid leaves of absence. Such a summary appears in the following tables. The study was sponsored by the Alberta Advisory Committee on Educational Research.

Collection of Data

Data were obtained as indicated:

1. A survey of the calendars of those Canadian universities offering degree programs in education.
2. A questionnaire to each Dean of Education in Canada regarding aid given in the form of assistantships, grants-in-aid, and special grants.
3. Letters to the Department of Education in each of the ten provinces respecting aid given to employees and to graduates in education generally.
4. Letters to Foundations and to Fellowship Committees listed in every available reference, asking for literature respecting their aid programs, or for confirmation of the information contained in the university calendars.
5. Letters to Teachers' and Trustees' organizations in each province respecting their programs of financial aid.
6. Questionnaires to school boards in Canadian cities having populations of 15,000 or over, respecting sabbatical and other forms of leave for educational study.

---

[1]First Canadian Conference of Deans of Education and Professors of Educational Administration and Supervision, **Report of Proceedings,** Toronto: CEA, p. 7.

7. Letters to service clubs mentioned in any university calendar respecting financial aid given for graduate study in education.
8. Letters of inquiry to selected businesses and industries.
9. Personal and telephone interviews.

TABLE I

POPULATION SAMPLED IN OBTAINING DATA ON FINANCIAL AID AND PERCENTAGE OF RESPONSE

Population	Number Contacted	Percentage of Replies
Universities	18	100%
Departments of Education	10	100%
Foundations, Fellowship Committees	36	100%
Teachers' Organizations	16	94%
Trustees' Organizations	17	88%
Urban School Boards[2] (15,000 population and over)	94	72%
Service Clubs	24	92%
Businesses	9	100%
Miscellaneous	15	87%
Totals	239	86%
Additional Data-Gathering Procedures:		
Personal and Telephone Interviews	31	.....
Follow-up Letters	50	.....

Deliminations of the Study

Field of study. At a number of universities teachers, or others about to become teachers, usually hold a B.A. degree or its equivalent before proceeding to the B.Ed. degree or to graduate work in education. Such persons might conceivably study in a wide variety of fields. In order to limit the fields of study in those areas not specifically designated for education, awards have been included only in those special fields that are usually taught in the senior high schools. These include Languages such as English, French, German, and Latin; and Humanities such as Art, History, Music, Literature, and Philosophy.[3]

In defining the fields of study that might be considered as open to students in education, most foreign languages, Classics, Social Work, Business courses, most of the Social Sciences, and most of the Natural Sciences have been excluded. It was felt that subjects of a

[2]Canadian Education Association, "Leave of Absence for Professional Study for Administrative and Supervisory Personnel", a report prepared for the CEA-Kellogg Project in Education Leadership, December, 1955.

[3]University of Toronto Calendar, Faculty of Arts, 1955-1956, p. 38.

specialized or technical nature require specialized training, and that students in them seldom enter the teaching field.

Financial limits. As it is felt that financial aid has to be fairly substantial in order to attract students into full-time graduate schools, awards of less than $100 have been omitted.

Bachelor of Education degree. In a number of Canadian universities the B.Ed. is considered a graduate degree. At all universities where the B.Ed. is awarded students may enroll for it after obtaining a degree such as the B.A. In order to obtain uniformity in this report, where a prior degree is listed as a requisite, students taking courses leading to a B.Ed. degree have been classified as doing graduate work, whether or not registered in the Graduate School.

Financial Awards

Space does not permit the inclusion of details respecting the awards offered by the individual Canadian universities. Only those awards offered by the University of Alberta are shown in full. A summary of the awards open to all Canadian students and tenable at any Canadian university is shown in Table II. Details respecting any of the awards may be found in appropriate university calendars.

UNIVERSITY OF ALBERTA

I. *Financial Awards Open to Canadian Students for Graduate Study at Any Canadian University*

A. *Scholarships and Fellowships*:
See Table II for summary of these awards.

B. *Bursaries and Loans*:
See Table II for summary of bursaries and loans.

TABLE II

FINANCIAL AWARDS OPEN TO CANADIAN STUDENTS FOR GRADUATE STUDY AT ANY CANADIAN UNIVERSITY

Name of Award	Value	Number	Conditions
Canadian Federation of University Women Junior Fellowship	$1,200	1	Open to women graduates not over 25 years old. Work must be approved.
Canadian Federation of University Women Margaret McWilliam Fellowship	$1,500	1	Open to women graduates not over 35 years old.
Canadian Federation of University Women Professional Scholarship	$1,000	1	Open to women graduates not over 35 years old. Library, education, professional schools.
Canadian Federation of University Women Travelling Fellowship	$1,500	1	Open to women graduates not over 35 years old.
Canadian Federation of University Women	$ 500	to 4	To assist former fellows to continue work previously started under CFUW.
Canadian Social Science Research Council Pre-Doctoral Fellowships	$2,000	about	For graduates doing post-graduate work.
Canadian Social Science Research Council Post-Doctoral Grants in Aid of Research	$ 800	not stated	For senior scholars for travel and research expenses.
Central Mortgage and Housing Corporation Senior Fellowships	Not fixed	to 2	For professional candidates for study of housing and urban development.
Central Mortgage and Housing Corporation Bursaries	$ 800	3	For graduates for studying housing and urban development.
Sir James Dunn Scholarships	$1,000	3	For graduates from elsewhere to study at University of New Brunswick.
Humanities Research Council of Canada Pre-Doctoral Fellowships	$2,000	to 12	For doing post-graduate work in Classics, Languages, Humanities.
Humanities Research Council of Canada Post-Doctoral Grant in Aid of Research	$ 500	Not stated	To assist senior scholars to meet expenses and travel costs.
Imperial Oil Graduate Research Fellowship	$1,250	1	Open to any graduate for research leading to a doctor's degree.

TABLE II—*Continued*

Name of Award	Value	Number	Conditions
Leonard Foundation Scholarships*	Not stated	Not stated	For British subjects, white race, Protestant religion. Leads to B.Ed.
Princess Alice Foundation Fund Scholarships	$1,500	Several	Preferably for women graduates, under 35 years of age.
Woodrow Wilson Fellowships	$2,250	150	Candidates nominated by University staff member.
BURSARIES AND LOANS			
Dominion-Provincial Student Loan Funds	$ 300		Loans available to full-time students in most degree courses. Available in P.E.I., N.S., N.B., Sask. as graduate loans. In P.Q., Man., Alta., B.C. for undergraduates only.
Dominion-Provincial Student-Aid Bursaries	$ 250 and $ 500		Type B bursaries for full-time Ont. university students.
D.V.A. and Armed Services Benefits			Under special circumstances benefits may be extended for graduate study.
I.O.D.E. Educational Foundation Revolving Loan Fund	$ 400		To assist any worthy young person. Not restricted to college education.

*Awards marked with a star (*) are for students studying for a B.Ed. degree.

II. *Financial Awards Open to Alberta Students for Graduate Study at Any Canadian University*

A. *Scholarships and Fellowships*:

1. John Walker Barnett Memorial Scholarship.
   Field of Study: Education.
   Value: $500.
   Number: One annually.
   Duration: One academic year.
   Conditions: Open to holders of a B.Ed. degree who are proceeding to graduate work. Holders must be bona fide residents of Alberta, and must register for intramural study in a graduate school within two years following the award.
   Tenable at: Any recognized university.
   Donor: Alberta Teachers' Association.
   Apply to: Alberta Teachers' Association, Barnett House, 9929 - 103 Street, Edmonton, by July 15.

2. Clarence Sansom Memorial Scholarship.

Field of Study:	Education.
Value:	$500.
Number:	One annually.
Duration:	One academic year.
Conditions:	Awarded to the winner of the Clarence Sansom Memorial Gold Medal in Education, for the highest general proficiency in the final year of the B.Ed. program. Payable upon registration for intramural graduate study, within two years following the award.
Tenable at:	Any recognized university.
Donor:	Alberta Teachers' Association.
Apply to:	Alberta Teachers' Association, Barnett House, 9929 - 103 Street, Edmonton, by July 15.

B. *Bursaries and Loans*:

1. Alberta Teachers' Association Professional Assistance Loans.
A fund has been set up by the Alberta Teachers' Association to provide loans to students in need of financial assistance to continue their preparation as teachers through intramural study. Priority for loans will be given to teachers in the field who have completed at least three years of professional education. Loans are interest free, but must be insured against loss due to death or disability. Apply to the Alberta Teachers' Association, Barnett House, 9929 - 103 Street, Edmonton, by July 15.

C. *Leaves of Absence*:

Employer	Length of Leave	Salary Provisions	Remarks
Department of Education	6 weeks	Full salary	
Calgary Public	1 year	$2,400	Sabbatical leave
Edmonton Public	1 year	$1,200	Sabbatical leave
Edmonton Separate	7 months	None	Individual cases. Some tuition, and expenses paid.
Lethbridge Div. No. 7	1 year	$1,200	Sabbatical leave
Lethbridge S.D. No. 51	1 year	$1,800	Sabbatical leave
Medicine Hat	6 weeks 1 year	50% salary $1,500 or 40%	Sabbatical leave
St. Mary's River	June 10 to 30	5 days' salary	Summer school semester plan.
Taber	1 year	$1,200	Sabbatical leave
Warner County	1 year	$1,200	Sabbatical leave

III. *Awards Tenable Only at the University of Alberta*

A. *Scholarships and Fellowships:*

1. Du Pont of Canada Scholarship.

   Field of Study: Secondary School Science Teaching.
   Value: $1,500 ($2,100 if a married man).
   Number: One annually.
   Duration: One academic year.
   Conditions: Awarded according to one of the following plans at the option of the University:
   (1) To a graduate or an undergraduate in an honours course in science or in a course in which science has been a major subject and who agrees to enroll for the following year for training as a secondary school science teacher.
   (2) Three summer school scholarships of $500 each to secondary school teachers for additional training in science. Recipients should be free to select any university conducting appropriate summer science programs.
   (3) To a science specialist already teaching in secondary school, whose ability would be improved by a year's postgraduate work in science.
   Donor: Du Pont of Canada Limited.
   Apply to: The Dean, Faculty of Education, University of Alberta.

2. John Macdonald and T. E. A. Stanley Scholarships.

   Field of Study: Education.
   Value: $500 each.
   Number: Two annually.
   Duration: One academic year.
   Conditions: Awarded to residents of Alberta who are graduates of other faculties and are proceeding to a B.Ed. program as a preparation for teaching.
   Donor: Alberta Teachers' Association.
   Apply to: Alberta Teachers' Association, Barnett House, 9929 - 103 Street, Edmonton, by July 15.

3. Stapells Graduate Scholarship.

   Field of Study: Language, Literature or Social Sciences.
   Value: $250.
   Number: One annually.
   Duration: One academic year.
   Conditions: Open to graduates of the University of Alberta who register in a full intramural program. Details of the proposed field of study must be given, together with confidential reports from three instructors.
   Donor: Mr. Frederick Stapells.
   Apply to: The Registrar, University of Alberta, by April 15.

4. Robert Tegler Research Scholarship.

   Field of Study: Unrestricted.
   Value: $900.
   Number: One annually.
   Duration: One academic year.

Conditions: Open to graduates of the University of Alberta. Details of the chosen field of study together with confidential reports from three instructors must be supplied.
Donor: The Robert Tegler Trust.
Apply to: The Registrar, University of Alberta, by April 15.

5. University of Alberta Research Scholarships.
Field of Study: Unrestricted.
Value: $900.
Number: Two; not always annually.
Duration: One academic year.
Conditions: Open to graduates of any Canadian university. Details of the chosen field of study together with confidential reports from three instructors must be supplied.
Donor: Board of Governors of the University of Alberta.
Apply to: The Registrar, University of Alberta, by April 15.

B. *Bursaries and Loans*:

1. Alberta Government Bursaries.
Field of Study: Education.
Value: $400 each, plus tuition fees.
Number: To a maximum of 40 annually.
Duration: One academic year.
Conditions: Granted to graduates of the University of Alberta or of other universities who enter the Faculty of Education at Alberta to obtain certification for teaching at the senior high school level. Candidates must undertake to teach at the high school level in the public schools of Alberta for two consecutive years following completion of the one year of training.
Donor: The Department of Education of Alberta.
Apply to: Coordinator of Teacher Education, Department of Education, Edmonton.

C. *Assistantships*:

Number	Kind	Weekly Work Required	Value
1	Research assistantship	10 hours	$ 520
1	Clinical assistantship	10 hours	$ 320
1,200 hours	Marking and reading assistantships	48 hours	$1,200

NOTE: Research grants are available in limited numbers and amounts through the Alberta Advisory Committee on Educational Research, University of Alberta.

Tabular Presentation of Data

In the tables that follow, the awards offered by the eighteen universities covered by this study are summarized. Table III shows the numbers and approximate values of scholarships, fellowships, and bursaries granted by various donors for graduate study at *any* Canadian university. In some instances awards are made in groups of no specified number or amount. Where this occurs, the letter "G" is used to indicate a group of an unspecified number and

amount. The table also differentiates between awards for strictly graduate study and those granted to holders of a previous degree who are studying for a B.Ed.

Table IV continues the classification in Table II by listing the numbers and approximate values of awards made by various donors for graduate study at the specific Canadian universities. In this table strictly graduate awards are indicated outside parentheses while awards for a B.Ed. program are within parentheses.

Table V shows all scholarships, fellowships, and bursaries available by universities and in the various fields of study. An examination of Table V indicates that in the whole of Canada a total of twenty-three single awards and three groups of awards are available to students in education doing strictly graduate work. It should be apparent that in each of the other fields of study, although education students may be free to apply for an award, they are in direct competition with other students, many of whom are specialists in the field in which the award is being offered. Those awards that are unrestricted as to field of study are open to the total university student population of Canada.

TABLE III

NUMBERS AND APPROXIMATE VALUES OF SCHOLARSHIPS, FELLOWSHIPS, AND BURSARIES GRANTED BY VARIOUS DONORS FOR GRADUATE STUDY AT ANY CANADIAN UNIVERSITY

Open to Students of	Universities	Departments of Government	Foundations and Fellowships	Teachers and Trustees	Service Groups	Businesses	Memorials and Persons
All Canada —(Grad)		2—G	32—$64,000 4—G		8—$7,200	4—$3,650 1—G	3—$3,000
(B.Ed.)							1—G
Newfoundland							
Prince Edward Island							
Nova Socta —(Grad)							
(B.Ed)		1—G					
New Brunswick —(Grad)		1—G					
(B.Ed.)							
Quebec —(Grad)				1—G			
(B.Ed.)		2—G					
Onatrio —(Grad)		4—$6,000					1—$3,000
	1—G						
(B.Ed.)		1—G					
Manitoba —(Grad)	2—$ 800				1—$ 150		
(B.Ed)							
Saskatchewan —(Grad)				1—$ 500	1—$ 500		
(B.Ed.)							
Alberta —(Grad)				2—$1,000			
(B.Ed)							
British Columbia—(Grad)	1—$ 200		1—$ 125		3—$ 450	1—$ 300	
(B.Ed.)							
TOTALS —(Grad)	3—$1,000	4—$6,000	33—$64,125	3—$1,500	13—$8,300	5—$3,950	4—$6,000
		4—G	4—G	1—G		1—G	
(B.Ed)		4—G					1—G

Table IV

NUMBERS AND APPROXIMATE VALUES OF SCHOLARSHIPS, FELLOWSHIPS, AND BURSARIES GRANTED BY VARIOUS DONORS FOR GRADUATE STUDY AT SPECIFIC UNIVERSITIES

Tenable at	Universities	Departments of Government	Foundations and Fellowships	Teachers and Trustees	Service Groups	Business	Memorials and Individuals	Miscellaneous
Memorial								
Acadia							3—$ 700	
Dalhousie	1—$ 500				1—$ 200 (1—$ 100)	1—$1,500	1—$ 100	
St. Francis X						1—$ 150		
Mt. Allison							(1—G)	
St. Joseph's								1—G
New Brunswick						1—$1,500	1—$ 750	
Bishop's								
Laval	1—G					1—$1,500		
McGill	2—G				8—$2,450 1—G		10—$3,950 2—G	
Montreal						1—$1,500		1—G
Ottawa								
Toronto	1—G					6—$8,000	7—$4,750 (2—$ 250) 1—G	
Manitoba		(6—$1,500)		(2—$ 400)	(1—$ 100)	1—$1,500		
Saskatchewan						1—$1,500	1—$ 500	
Alberta	2—$1,800	(40—$24,000)		(2—$1,000)		1—$1,500	2—$1,150	

British Columbia			2—$1,000	(5—$ 700)	1—$ 100 (7—$ 700)	1—$1,500 1—G	4—$ 650	
TOTALS	3—$2,300 4—G	(46—$25,500)	2—$1,000	(9—$2,100)	10—$2,750 (9—$ 900) 1—G	17—$23,150 1—G	29—$12,550 (2—$ 250) 3—G(1—G)	2—G

NOTES: 1: The letter "G" indicates a group of awards.
2. Numbers outside parentheses indicate awards for graduate degree programs.
3. Numbers inside parentheses indicate awards for B.Ed. program only.
4. Amounts shown are only approximate. If awards are stated as the interest from a fund, estimates are made at 3%. Values are not shown for groups of awards, as the numbers awarded are not usually stated.

TABLE V

SCHOLARSHIPS, FELLOWSHIPS, AND BURSARIES AVAILABLE IN VARIOUS FIELDS OF STUDY AT CANADIAN UNIVERSITIES

Tenable at	Unrestricted	Several Fields	Education		Humanities	Languages	Others
			Grad.	B.Ed.			
Any University	3 3 (Man.) 1 (Sask.) 6 (B.C.)	17 4—G	G (N.B.) G (P.Q.) G (Ont.) 5 (Ont.) 1 (Sask.) 2 (Alta.)	1—G G (Ont.)	1		25 2—G
Several Universities	3			G (N.S.) 2G(P..Q.)			
Memorial							
Acadia		1			1	1	
Dalhousie	2		1	1		1	
St. Francis					1		
Mt. Allison				1—G			
St. Joseph's		1—G					
New Brunswick			1		1		
Bishop's							
Lavel	1—G		1	2—G			
McGill	14 4—G	3 1—G	2			1	
Montreal		1—G	1				
Ottawa							
Toronto	5 1—G	1 1—G	5	2	2		
Manitoba			1	9			
Saskatchewan		1	1				
Alberta	3	1	1	42			
British Columbia		2 1—G	1	12	2		3
TOTALS	40 6—G	26 9—G	23 3—G	66 8—G	8	3	28 2—G

NOTES: 1. The letter "G" indicates a group of awards.

2. Restricted awards have the name of the province in parentheses in which the award is available.

Table VI

SCHOLARSHIPS, FELLOWSHIPS AND BURSARIES AVAILABLE TO STUDENTS HOLDING AN UNDERGRADUATE DEGREE AND PROCEEDING TO A BACHELOR OF EDUCATION DEGREE

Donor and University Where Tenable	Number of Awards	Value
Leonard Foundation (Tentable at most universities)	Several	Not stated
I.O.D.E. (Dalhousie)	1	$100
Department of Education, Nova Scotia (Five universities)	All Students	Fees
Herbert S. Sharp (Mount Allison)	Several	$500 each
Department of Health and Social Welfare, Quebec Provincial Government of Quebec (Most universities)	Several Several	$400 each $1,200 each
Hugh W. Bryan, and John T. Crawford (Toronto)	2	$250
Department of Education, Ontario (Any university)	Several	$250 and $500
Department of Education, Manitoba Manitoba Teachers' Society I.O.D.E. (Manitoba)	6 2 1	$1,500 $ 400 $ 100
Alberta Teachers' Association* Department of Education, Alberta (Alberta)	2 to 40	$1,000 $400 plus fees
B.C. Parent-Teacher Federation B.C. Teachers' Federation Groups I.O.D.E. Chapters National Council of Jewish Women (British Columbia)	1 5 5 1	$100 $700 $500 $100

*Macdonald and Stanley scholarships.

TABLE VII

ASSISTANTSHIPS AND SPECIAL GRANTS AVAILABLE FOR GRADUATE STUDY IN EDUCATION AT CANADIAN UNIVERSITIES

Granted by	Assistantships		Grants	
	Number	Value	Number	Value
Memorial				
Acadia	4	$ 950		
Dalhousie	4	$1,500		
St. Francis				
Mt. Allison	3	$ 400	Many + 2	Fees in Ed. $ 228
St. Joseph's			4	$ 600
New Brunswick	3	to $800		
Bishop's				
Laval	2	$2,000	2	$ 425
McGill				
Montreal			2	$1,000
Ottawa	2	$2,000	2	$ 200
Toronto	6	to $7,900	1	$ 600
Manitoba	1	$1,000	1	$ 50
Saskatchewan	1	$ 450		
Alberta	Several + 2	to $1,200 $ 840		
British Columbia	2	$1,650		
TOTALS	30 + Several	to $20,690	14 + Many	$3,103 + Fees in Ed.

TABLE VIII

SCHOOL BOARDS GRANTING LEAVES OF ABSENCE FOR FURTHER EDUCATIONAL STUDY IN CANADIAN CITIES OF 15,000 POPULATION AND OVER

Number of Boards Reporting	Grant Leave Up To 6 Weeks		Grant Leave For 7 Months to 1 Year			Other Financial Aid Given	Previous Service Required	Considering Granting Leave	Grant No Leave
	Without Salary	Full Salary	Without Salary	Full Salary	Part Salary				
8 in Atlantic Provinces		4	6	1		1	1—On merits 1—1 to 4 yr. 6—Not stated		
11 in Quebec	1	4	5	1		3	2—On merits 2—1 to 4 yr. 7—Not stated		4
31 in Ontario	10	8	12	1	9	4	4—On merits 2—1 to 4 yr. 4—5 yr. & over 21—Not stated	2	8
18 in Western Provinces	2	5	11	1	12		2—to to 4 yr. 11—5 yr. & over 5—Not stated	1	
TOTALS	13	21	34	4	21	8	7—On merits 7—1 to 4 yr. 15—5 yr. & over 39—Not stated	3	12

Table VI elaborates upon the awards granted to students who already hold one undergraduate degree and are proceeding toward a B.Ed. degree. This information is also contained in the B.Ed. column in Table V.

In the original report two tables were included dealing with loans available for graduate study at the Canadian Universities. Space does not permit their inclusion here. Table VII shows the extent to which assistantships and special grants are made available for graduate study. In a number of instances the amounts or numbers are flexible to meet the requirements of the university department concerned.

Table VIII shows the number of school boards in Canadian cities of 15,000 population and over that grant leaves of absence to their employees for educational study. Such leaves might be used for undergaduate study, for study abroad, and for similar purposes as well as for graduate study.

No attempt was made in the report to the Second Canadian Conference to draw conclusions from the data presented. The purpose of the report was to present the information so that action by the universities might be soundly based.

BIBLIOGRAPHY

Canadian Education Association. *Leave of Absence for Professional Study for Administrative and Supervisory Personnel.* Report prepared for the CEA-Kellogg Project in Educational Leadership, December, 1955. Toronto: Canadian Education Association, 1955. Mimeographed.

First Canadian Conference of Deans of Education and Professors of Educational Administration and Supervision. *Report of Proceedings.* Toronto: Canadian Education Association, 1955. Mimeographed.

University of Toronto Calendar. *Faculty of Arts, 1955-1956.* Adapted from p. 38.

TO TEACH, OR NOT TO TEACH

J. D. Aikenhead
Faculty of Education
University of Alberta

Why do some young people choose to teach, and some do not? Why do teachers stay teaching? What reasons impel some teachers to leave the classroom, try out business, and later return to teach? What influences cause men and women holding a permanent teacher's certificate to leave the profession?

These questions about teacher recruitment and retention are basic to a solution of the teacher shortage in Alberta. The estimated shortage for Alberta when schools opens in September 1956, is 1,243, and for 1960 an aggregate of 6,382.[1] A Canadian educator stresses the quality factor when he writes that 1,000 of the 4,000 training to teach each year would be screened if there was no teacher shortage.[2]

Other studies have been made in Alberta on the problems of recruitment. Clarke and Pilkington reported on reasons given by education students for selecting teaching as a career.[3] Melsness gave findings on a more limited sample.[4] This paper presents the present writer's research covering both problems. Part I is a summary of a doctor's dissertation[5] on teacher recruitment in the four western provinces. Part II deals with teacher retention in Alberta.

1. Recruiting Teachers

This research aimed to find out why young people in the four western provinces choose to teach, or not to teach. Young people who were studying in the last two years of high school or the first two years past high school completed 3,807 questionnaires, a return of over 82 per cent. To insure an adequate cross-section assistance was given by principals, school superintendents, officials of the Departments of Education, university officials, the General Secretary of the British Columbia Teachers' Federation, as well as many others. A grant from the General Research Council of the University of Alberta assisted.

1**Facts About Education's Number One Problem**, Conference Report (December, 1955), p. 6.

2**Yearbook of Education**, p. 233.

3S. C. T. Clarke and W. Pilkington, "Why Teaching is Chosen as a Career" **The Alberta Journal of Educational Research** (March, 1955), pp. 44-52.

4H. C. Melsness, "More Information, Please" **The Alberta School Trustee** (May, 1953), pp. 8-9.

5**J. D. Aikenhead, To Teach, or Not To Teach.**

The dissertation includes 113 tables with as many pages of supporting narrative. Since excerpts would make this report unduly long, the established conclusions are listed.

1. Young people stated they avoided teaching mostly because of monotony, slow promotions, few highly paid top positions, and low salaries.

2. The teacher shortage is greatest where the salaries are the lowest.

3. The shortage is greatest where there are fewer larger units of school administration, or where these units have not been operating for long periods.

4. Young people who do not plan to teach were not deterred to a marked degree by an alleged lack of prestige, critical parents, rude children, or the hard work of the classroom.

5. Almost no students were prevented from finishing high school, or entering college, by the cost or difficulty of finishing high school.

6. Students chose to teach mainly because they liked children and young people, security, teaching hours and holidays. Next in order of importance were doing good for others, the prestige of the teaching profession, a satisfactory salary and adequate sick benefits with pension.

7. The opportunities afforded by teaching for travel and self-improvement did not influence many students.

8. Bursaries tend to increase the numbers training to teach.

9. Small rural high schools send the largest proportion of young people to train for teaching.

10. Pamphlets from Departments of Education are influential in persuading young people to enter the teaching profession.

11. Most young people decide to teach when sixteen to nineteen years of age, but a significant number decide shortly before or after.

12. Teachers come from homes where the level of formal schooling is below that of their peers who do not plan to teach.

13. The basic two-year training for a permanent certificate tends to encourage students to plan for two year's training.

14. Students are not always given accurate information on teachers' salaries.

15. High school students might have more opportunities to teach in the lower grades in their school under capable teachers.[6]

These findings show why young people in Western Canada choose to teach or to avoid teaching. However, keeping good teachers in the classroom is the more vital problem.

[6]Ibid., pp. 179-180.

2. Retaining Teachers

According to a report of the National Conference of Canadian Universities the "most difficult single problem that must be solved . . . is the teacher shortage."[7] Both Stinnet[8] and the N.E.A.[9] give data which are alarming. Dr. Garnet Page of the Chemcial Insitute of Canada, shows that the classrooms must have science teachers to train persons for industry.[10] Page also points out how Russia sends her surplus scientists, fully indoctrinated, to undeveloped countries. In Russia almost four times as many young people are training at university level as in Canada, when the populations are compared.[11] An Alberta study was completed by Murray on annoyances which cause teachers to leave the profession.[12]

Teachers, Returned-teachers, and Ex-teachers

The present study was made possible with the help of Mr. W. E. Frame, Chief Superintendent of Schools, Government of Alberta, that of several school superintendents, and a grant from the General Research Council of the University of Alberta. A total of 120 persons were interviewed from Peace River to Lethbridge, St. Paul, Oyen, Brooks to Banff. All persons were adults who held permanent teacher's certificates. The 120 persons were equally divided into three groups. Forty were teachers with permanent certificates, regularly employed since they began their teaching careers. This group will be referred to as "teachers". Another 40 teachers holding permanent certificates had left the classroom for a full year or more to earn their living at some other vocation before deciding to return to teach. This group is termed "returned-teachers". (Married women who left to keep house, or women who left to marry were not interviewed.) The third group of 40 ex-teachers appear to have left the profession permanently. (Again, married women who left for housekeeping duties, or women who left to marry, were not included in the study.)

Marital Status: Dependents

More married men returned to the classroom after trying some other work. This one would expect. Single men were slower to return to teach, as indicated in Table I. The reasons given during

---

7The National Conference of Canadian Universities, Proceedings, 1955, p. 63.

8T. M. Stinnet, "The Teacher Supply," **The School Executive** (January, 1955), p. 69.

9"Teacher Supply and Demand Report," **Journal of Teacher Education** (March, 1955), pp. 23-71.

10Garnet T. Page, "Science and Mathematics Teaching in our Secondary Schools," **Chemistry in Canada** (March, 1956), pp. 35-37.

11Garnet T. Page, "Education and Professional Manpower in the U.S.S.R.," **Chemistry in Canada** (May, 1956), pp. 35-39.

12Thomas H. Murray, **An Investigation into the Annoyances and Frustrations Which Cause Alberta Teachers to Quit Teaching, The Alberta Journal of Educational Research (September, 1955), pp. 17-23.**

the interviews by all the men were to the effect that the work was not as lucrative as promised, that the breadwinner was away from home too much, that working conditions affected health adversely, that pressure to achieve quotas was ruining health, and that most often the standard of living was lower. A few men admitted, "I wasn't cut out for business and it cost me a lot to find out."

It would seem that among the ex-teachers married women do not likely leave the profession to accept other employment. (See Tables V to VIII.)

Men have more dependents than women and women with dependents seem to favor the classroom. Table II shows that there were no married women among the ex-teachers. The interviews gave more information. Several men mentioned they would have stayed out of the classroom if they hadn't a family. As one man said, "If you are married you can only be out of work so long."

TABLE I

MARITAL STATUS OF PERSONS INTERVIEWED

Status	Teachers		Returned-teachers		Ex-teachers	
	Men	Women	Men	Women	Men	Women
Married	10	7	24	4	29	.....
Widowed	.....	1	1	1	2	.....
Divorced	.....	.....	1	1	1	.....
Single	5	17	3	5	1	7

Professional and Academic Training

Table III shows that most women had the minimum academic and professional training. Those women actually teaching regularly had more professional training and a superior background when compared with the women in the returned-teacher and ex-teacher groups. The numbers interviewed were too meagre to allow a careful comparison between the men in the teacher-group and the returned teacher groups. The number of men holding degrees is significant in the ex-teacher group. Three held one professional degree, four one academic degree, two had earned two degrees, one held one professional and one academic degree, and one ex-teacher held two professional and one academic degree.

TABLE II

NUMBER OF DEPENDENTS OF PERSONS INTERVIEWED

	Teachers									
	Men					Women				
	Number of Dependents					Number of Dependents				
	0	1	2	3	4 plus	0	1	2	3	4 plus
Married	...	2	1	5	2	5	1	1	...	...
Widowed	...	...	...	...	...	...	1	...	...	...
Single	4	1	...	...	...	16	1	...	...	...
	Returned-teachers									
Married	1	3	4	7	9	2	1	...	...	...
Widowed	...	...	...	1	...	2	...	...	...	...
Divorced	1	...	...	...	...	...	...	1	...	...
Single	3	...	...	...	...	4	1	...	...	...
	Ex-teachers									
Married	1	2	8	12	6	...	...	...	...	...
Widowed	1	1	...	...	...	...	...	...	...	...
Divorced	1	...	...	...	...	...	...	...	...	...
Single	1	...	...	...	...	7	...	...	...	...

Teaching and Business Experience

Table IV shows most teachers were without business experience. The returned-teachers showed a wide range within, and outside, the classroom. In the ex-teacher group about half of the men had between six and ten years of teaching experience, and two thirds of the sample had been in business between six and fifteen years. Of the seven women interviewed four had taught more than five years, and one had taught 25 years. This women left the profession in good standing, of her own volition, with credits past a bachelor's degree. Data such as this show the need for more time for research broad enough to find why so many good teachers leave when they are such an asset to the children in Alberta.

Respondents' Opinions

Twelve questions were asked covering reasons why teachers

teach, and 14 reasons were proposed to persons who had left the classroom. The reasons are shown in Tables V to VIII. As mentioned, the questionnaire was the open-end type and comments were noted as the interview progressed, and later recorded. Each person was asked to give the extent of influence for each reason. For example, when a returned-teacher was asked by the present writer, "Did you return to teach because teaching offers security?" and gave the answer as "Yes", the interviewer then determined where this reason could be located on a five-point scale. That is, in this case the returned-teacher was asked how much this reason influenced him to return to the classroom. Was it the deciding factor (4 points), or very strong (3 points), or definite (2 points), or little (1 point), or no influence (0 points). Later all the responses of the 40 persons interviewed in each group were taken in the five-point extent-of-influence scale and multiplied by 4, 3, 2, 1, and 0 respectively. These products were added to secure a "weight-total" for each reason for men and for women. This weighted-total not only shows the extent of the influence of each reason, but allows a comparison that is somewhat quantitative.

TABLE III

YEARS OF PROFESSIONAL AND ACADEMIC TRAINING OF PERSONS INTERVIEWED

Years professional training	Teachers		Returned-teachers		Ex-teachers	
	Men	Women	Men	Women	Men	Women
1	4	18	13	7	25	7
2	3	4	6	3	1	.....
3	2	1	5	.....	1	.....
4	6	1	3	1	4	.....
5	.....	1	1	.....	2	.....
7	.....	.....	1	.....	.....	.....
Years academic training						
Gr. XI	.....	.....	.....	.....	1	.....
XII	7	20	19	11	22	7
plus 1	1	2	1	.....	1	.....
2	.....	1	3	.....	2	.....
3	2	.....	2	.....	5	.....
4	4	2	4	.....	.....	.....
5	1	.....	.....	.....	2	.....

TABLE IV

YEARS OF TEACHING AND BUSINESS EXPERIENCE OF PERSONS INTERVIEWED

Years teaching experience	Teachers		Returned-teachers		Ex-teachers	
	Men	Women	Men	Women	Men	Women
2	.....	1	.....	.....	2	.....
3	.....	2	1	.....	2	1
4	4	4	4	.....	3	2
5	.....	4	1	2	2	.....
6 to 10	2	4	9	5	16	2
11 to 15	1	6	4	3	5	1
16, plus	8	4	10	1	3	1
Years business experience						
0	11	22	.....	.....	.....	.....
1	1	.....	5	3	.....	.....
2	.....	2	4	5	1	2
3	1	.....	3	1	1	1
4 to 5	1	1	5	1	5	1
6 to 10	.....	.....	8	1	10	2
11 to 15	1	.....	2	.....	13	1
16, plus	.....	.....	2	.....	3	.....

Reasons Given for Teaching

Men teach mainly because they like children and young people, as shown in Table V. All the other reasons are ranked for the men respondents. Almost half the men remarked, "I like to teach. That's it." Two other reasons have equal weight, altruism and a liking for colleagues. Holidays, hours and salary appear to be only half as persuasive as a liking for youngsters.

Women teach too because they like their pupils and students as a first reason. However, this ranking for the men does not coincide with that for the women. For women the cluster of reasons which covers security aggregates 189, while the total for men is 66. Even with the larger number of women the figure is significant. Women give a wider variety of responses on the extent of influence for each reason. Men tended to mark either extreme on the five-point scale.

Reasons for Leaving Given by Returned-teachers

The reasons given by these forty persons, ranked for the 29 men, are listed in Table VI. As mentioned, women who left to marry were not interviewed. The table shows that for the men low salaries

not only lead the list but were almost four times as important as slow promotions, which rank second. Only two other reasons seem to influence men strongly: not many highly paid top positions, and the subjective nature of teaching. The weighted totals show that financial reasons far outweigh other influences: 143 to 72 respectively. For women the reasons are not ranked as listed in the table, though salaries were judged most important. Women gave as the next most persuasive reasons the monotony of teaching, large classes, and substandard living conditions in some parts of the country. The other reasons listed in the table were not considered seriously by the women. One woman stated:

> An itinerant family caused a row to embroil the whole community. I got out at the end of June to take work in another town in business. I was sure the trouble making family would leave and I could get my old job back, but if I don't, there were lots of others.

One man explained:

> I was given an attractive offer selling, a 23 per cent increase in salary, expenses, and a new car to drive, week-ends too for the family, with a commission over the salary. My wife and I decided to try it. I could

TABLE V

REASONS GIVEN FOR TEACHING: TEACHER GROUP

Reason	Weighted total	
	Men	Women
I like children and young people	42	77
Teaching offers security	32	62
I like to teach. The work has an absorbing appeal for me	28	20
I feel I do good to other people as a teacher	26	45
I like the teachers I know	26	32
The holidays appeal to me	21	34
I like the hours	20	41
The salary satisfies me	18	38
I believe teaching has prestige	13	20
Pensions and sick benefits seem satisfactory	8	22
A permanent teacher's certificate is valuable to a man, who can leave teaching to take a chance in business, and always return to teaching as a sure means of earning a living . . . . the same is true for a woman; and further she may marry and teach, or teach as a widow	8	67
I teach because I cannot do heavy physical labor	.....	6

TABLE VI

REASONS GIVEN FOR LEAVING TEACHING: RETURNED-TEACHERS

Reason	Weighted total	
	Men	Women
Salaries for teachers were too low	92	26
Promotions were slow in teaching	25	1
There are not many highly paid "top" positions	19	1
I could not prove my efficiency as a teacher	16	0
Living conditions in the country did not suit me	10	8
Personal freedom was often restricted on hours for leisure, friends, types of amusement, place to live, joining certain groups, pressure to overload me with community chores that would not ordinarily be taken by one person	10	2
Teaching did not have prestige	10	0
Classes were too large to allow a teacher to do good work	9	8
Teaching was too monotonous for me	9	15
Pensions were low after long service	7	3
Pupils' parents criticised teachers too much	5	0
A few children were rude and I did not like this	2	0
Teaching was too hard work	2	0
A large percentage of teachers never marry	0	1

always come back to teach. After a year and a half the firm started to cut my expenses. I said I'd quit. They restored the cuts. I started to look for my old job, but that city wouldn't take a teacher back on who left for business. I got this job in the country and like it. I was in business for two years. I saved a good stake.

These 29 men and 11 women changed their minds and returned to teach. Table VII shows that they wanted steady employment. When weighted totals are considered, 254 is the figure for men and 80 for the women. In consideration of the number of men and women, these totals show about the same emphasis. It is interesting to note that in business some men found that teaching had prestige. One returned man said, "I never mixed much with business people when I taught. In business I had to, and found the better people respected teachers." Teachers were also persuaded to return to

the classroom because they like children and their peers, and enjoyed freedom in hours of work. Altruistic motives were important as well, though these had less influence than steady employment. More women than men returned to the classroom because they liked to teach.

TABLE VII
REASONS FOR RETURNING TO TEACH:
RETURNED-TEACHERS

Reason	Weighted total	
	Men	Women
A permanent teacher's certificate is valuable to a man, who can leave teaching and always return as a sure means of earning a living . . . . the same holds true for a woman	88	29
Teaching offers security	87	28
I like children and young people	54	26
I feel I do good to other people as a teacher	47	27
I like the teachers I know	44	21
The holidays appeal to me	42	24
I like the hours	38	18
The salary satisfies me	37	15
Pensions and sick benefits seem satisfactory	30	8
I believe teaching has prestige	29	17
I like to teach. The work has an absorbing appeal for me	24	16
I teach because I cannot do heavy physical labor	12	0

Ex-teachers

Table VIII shows definitely that the 33 men and the seven women interviewed did not leave the profession for the same reasons. The influences given by the men are again ranked. Salaries were not only twice as important as any other reason, but the weighted totals of salary, the lack of highly paid top positions, and slow promotions are the only reasons which carry a significant weighted total. These aggregate 165 for the men, and proportionately for the women, 35. Women left mainly because living conditions in the country did not suit them, large classes were a burden, and the social life was restricted. A young women related:

> I came to Alberta from Saskatchewan. I was experienced and judged to be capable. I hired by 'phone. The superintendent said he'd put me in a

close-in school, but I found it far-out. The roads were terrible. The superintendent came first in October. He was plastered, and had left his car in a mud-hole. He was the first man I'd seen wearing a tie since Labor Day. There was only one place to board. The husband and wife never spoke. They didn't seem to dislike each other, but they did dislike talking. The district was divided in religion. Neither side believed it was right to enjoy yourself. It was a quiet year for everybody. I had to get out in June.

The heavy teaching load caused this response:

You can't do justice to 25 kids in a country school with all grades. I tried a town school next, one grade with 45. I nearly went crazy. At University they taught us to be on the playground with enthusiasm. The other teachers resented me being better liked for this. There were a lot of chores to do. We had good staff meetings, but I saw how poorly I was doing in testing, in teaching, and almost nothing in community work, and no social life at all. I gave up.

TABLE VIII

REASONS FOR LEAVING TEACHING: EX-TEACHERS

Reason	Weighted total	
	Men	Women
Salaries for teachers were too low	100	3
There were not many highly paid "top" positions	44	2
Promotions were slow in teaching	21	1
Classes were too large to allow a teacher to do good work	14	17
Teaching did not have prestige	13	2
Pensions were low after long service	12	1
Teaching was too monotonus for me	12	6
I could not prove my efficiency as a teacher	11	6
Pupils' parents criticized teachers too much	9	8
Living conditions in the country did not suit me	9	16
Personal freedom was often restricted on hours for leisure, friends, types of amusement, place to live, joining certain groups pressure to overload me with community chores that would not ordinarily be carried by on person	6	7
Teaching was too hard work	6	7
A few children were rude and I did not like this	4	5
A large percentage of teachers never marry	0	12

Another teacher stressed ethics: "A few incompetent teachers gave the rest a bad name, and the weak teacher is not disciplined by the strong ones on most staffs." Still another said: "I did not want to be hard up always. Experienced teachers with degrees and getting more were always short of money, and of time to play with their kids." One man, judged by his peers to be able, explained:

> I left with genuine regret. I had two and a half years of Varsity, but felt I would need to invest another $3,000. and spend little time with my family for about four years. I doubled my salary when I started in business, and it is now twice what the highest paid principals receive.

Conclusions

Part II of this report shows that the teacher-group feels secure and happy in teaching, but they are not satisfied with the pay, pensions and prestige. The men in the returned-teacher group claim that monetary considerations and the subjective aspects of teaching influenced most men to leave. The women left because of the monotony of the classroom, large classes and substandard life in the country. This group returned to teach mainly because business did not offer what they had expected, while teaching assured a steady income, young people to help, and congenial colleagues. In the ex-teacher group the men left mainly because they felt that each community did not think enough of their services to pay what they later found they could earn in business.

The local people could have alleviated some of the annoying conditions not related to money. Their decision caused the profession to lose eleven men with one or more university degrees and successful experience as teachers and principals. The women in the ex-teacher group decided that living conditions in the country were not pleasant and that large classes were frustrating.

Proposals

Two writers[13] in the United States urge: (1) non-professional and sub-professional assistants for teachers, (2) professional assistants and internes, (3) more economical use of audio-visual aids, (4) delegation to the community of some areas now covered in school, (5) more dignity in the teaching profession, (6) parental recognition of the good services teachers give, and (7) improved public relations.

The swelling enrolments in schools must have teachers. One extreme plan is to lower standards for teaching and pupil performance, and to have large classes. This plan will fail as it has since Lancaster and Bell devised their scheme in the last century. Another extreme plan is to invest money and imagination to meet the

[13]W. H. Conley, "Creative Look at the Teacher Shortage," **The American School Board Journal** (April, 1955), p. 60; and R. C. Woellner, "Our Contributions to Overcoming Teacher Shortages," **The School Review** (September, 1954), p. 317.

crisis in Canadian education. Money and imagination have never been invested at the same time in Canada on a scale large enough to shown parents and business men what quality can be worth in education. The teacher shortage must be met if Canadian children are to have an education appropriate to Canada's resources today and her responsibilities tomorrow.

BIBLIOGRAPHY

Aikenhead, J. D. *To Teach, or Not To Teach.* Ed.D. Dissertation, University of Oregon, 1954.

The Alberta Home and School Association. *Facts About Education's Number One Problem* (Special Conference Report). The Association, Calgary, 134 - 8th Ave. East, December 3, 1955.

Clarke, S. C. T. and Pilkington, W. "Why Teaching is Chosen as a Career." *The Alberta Journal of Educational Research* 1:1 (March, 1955), pp. 44-52.

Conley, W. H. "Creative Look at the Teacher Shortage." *American School Board Journal* 130:60 (April, 1955), p. 60.

Melsness, H. C. "More Information, Please." *The Alberta School Trustee* XIII:6 (May, 1953), pp. 8-9.

Murray, Thomas H. "An Investigation into the Annoyances and Frustrations Which Cause Alberta Teachers to Quit Teaching." *The Alberta Journal of Educational Research* 1:3 (September, 1955), pp. 17-33.

The National Conference of Canadian Universities. *Proceedings.* The University of Toronto, 1955.

Page, Garnet, T. "Science and Mathematics Teaching in Our Secondary Schools." *Chemistry in Canada* VIII:3 (March, 1956), pp. 35-37.

Page, Garnet T. "Education and Professional Manpower in the U.S.S.R." *Chemistry in Canada* VIII:3 (May, 1956), pp. 35-39.

Rivers, F. S. and Jackson, R. W. B. "Teacher Supply in Canada." *Canadian Education* VIII:3 (June, 1953), pp. 3-21.

Sheffield, E. F. *Canadian University and College Enrolment Projected to 1956.* Ottawa, Dominion Bureau of Statistics, 1955.

Stinnet, T. M. "The Teacher Supply." *The School Executive* 74:1 (January, 1955), p. 69.

"Teacher Supply and Demand Report." *Journal of Teacher Education* VI:1 (March, 1955), pp. 23-71.

Woellner, R. C. "Our Contributions to Overcoming Teacher Shortages." *The School Review* 62 (September, 1954), pp. 315-319.

Yearbook of Education, London, Evans Bros., 1953.

AN EVALUATION OF THE EFFECTIVENESS OF A STANDARDIZED TEST OF LANGUAGE

C. H. Sangster
Supervisor of Elementary Instruction
Medicine Hat

The Problem

In an analysis of the strengths and weaknesses of an individual or groups of individuals, the effectiveness of measurement depends primarily on the reliability and validity of the instrument employed. When a test is to serve as a diagnostic instrument as well as a test of general achievement, it must yield component as well as total scores, each of which must be reliable and valid.

Test reliability and validity depend not only upon the content in general, but also upon the reliability and validity of the individual test items. No amount of statistical work can make up for weak items.

In selecting items for a test a number of considerations must be taken into account. One of these is the level of difficulty. What constitutes the optimum level of item difficulty in an achievement test has not been agreed upon among test authors. Some prefer approximately equal numbers of items at all levels of difficulty; others prefer to use a few easy and a few difficult items with the majority near the 50 per cent level of difficulty. There appears to be general agreement, however, that "the average difficulty of all items should be about 50 per cent."[1]

Another factor is the discriminating power of the test item. An item is said to have maximum discriminating power if every pupil who responds correctly ranks higher in achievement than any pupil who fails on the item. An item has zero discriminating power when there is no difference in achievement of students who succeed on an item and those who fail.

Statistical techniques for expressing the discrimination index of an item provide a measure of the relationship between success on the item and an independent criterion or the total score. A low index may be the result of technical weaknesses in the composition of a test item. Ambiguities, irrelevent clues and "catch" questions may make items ineffective. Low indices may also indicate that the "items are measuring rather specific points and are not highly

[1]H. E. Hawkes and others, **The Construction and Use of Achievement Examinations, p. 32.**

correlated."[2] If the area being tested includes a type of achievement that is not highly related to other achievements within the same area, then items testing for that type of achievement would not be expected to yield high discrimination indices.

This immediately establishes a fundamental difference in the content of general achievement and diagnostic tests. A general achievement test is one designed to express in terms of a single score a pupil's achievement relative to others in a class. The meaningfulness of a single score as a measure of general achievement depends upon the homogeneity of the field being tested. If low negative correlations exist between specific types of achievement, then no single measure can adequately describe a pupil's achievement in all these types simultaneously. Since the significance of a single score in a general achievement test depends upon the discriminating power of its items, it is desirable that their selection be restricted to those which correlate most highly with the total score so that a maximum in test homogeneity is achieved.

In a diagnosic test, items are selected to provide unitary subtests which measure independent areas of difficulty throughout the subject in question. Correlations between items and the total score, under such conditions, are frequently low or negative. Consequently, composite or total scores have little or no meaning.

It is not possible to have it both ways. The high correlation between items and total score, which makes for good measurement of general achievement, interferes with effective diagnosis. Likewise, the lack of correlation between component scores in diagnostic tests weakens effective measurement of general achievement.

Related Studies

As the reliability and validity of a test depend upon the characteristics of the items, it follows that a test may be improved through the selection, substitution or revision of its items. With the growing refinement of test construction, the use of item analysis to improve test efficiency has received increasing attention.

A study conducted by John E. Anderson[3] illustrates the bearing of item analysis upon the discriminative power of an English achievement test.

Following the administration and marking of an English examination, the discriminative power of each item was determined. Of a total of 222 items, eighty-six were selected as "good," eighty-three as "poor," and fifty-three items were found neutral. The tests were

[2]**Fredrick B. Davis, Item-Analysis Data, p. 3.**

[3]John E. Anderson, "The Effect of Item Analysis upon the Discriminative Power of an Examination," **Journal of Applied Psychology** (Vol. 19), pp. 237-244.

rescored to obtain the total score for each pupil on the good and the poor items. An analysis of the scores produced the following results.

1. The scores from the good items correlated .95 with the total test score.
2. The scores from the poor items correlated .45 with the total test score.
3. The correlation between scores on the good and poor items was .21.

Anderson says "the fact that the reduced examination composed of good items, constituting approximately one-third of the total number of questions, correlates .95 with the total examination score, shows that the good items alone are almost as valuable as the total examination."[4]

To assess the validity of the full test, of the good items and of the poor items, scores were correlated with final English grades, gradings on three book reports, and scores on an Iowa English test. Coefficients of reliability were computed by the split-half method.

Anderson concludes as follows:

1. The score on the good items predicts the total test score with high accuracy and predicts the final grade in English, book report grades and the Iowa English score better than does the total test.
2. Item analysis resulted in the production of a good item test which had almost as great reliability as the total test and which, if it had possessed the same number of items, would have had noticeably greater reliability.

A similar study was conducted by C. H. Lawshe, Jr. and James S. Mayer[5] in which they compared two methods of item analysis to determine which produced the more reliable test. They also compared the reliability coefficients of a short test composed of good items with that of a longer test composed of good and poor items.

Two indices of discrimination were determined for each of 300 items by the correlation method devised by Flanagan and by the D-value method based on Lawshe's nomograph. The best 20, 40, 60 and 100 items were selected from the original 300 by the use of the two item analysis methods.

The reliabilities of the tests were computed by the split-half method and the following comparisons were made.

1. In selecting forty and sixty items, the two methods produced tests with no difference in reliability.

[4]Ibid., p. 238.

[5]C. H. Lawshe and others, "The Effect of Two Methods of Item Validation on Test Reliability," **Journal of Applied Psychology** (Vol. 31), pp. 271-277.

2. The Flanagan method produced the more reliable test of twenty items and the D-value method yielded the more reliable test of one hundred items.
3. The difference between the reliability of the test produced by the Flanagan method and the original test was not significant. The D-value test was more reliable than the original 300 item test.

These two studies indicate that through the use of item analysis, it is frequently possible to shorten a test and at the same time to increase its reliability.

Research Design

The present study was designed primarily to determine the effectiveness of the California Language Test items as general achievement test items for selected groups of Alberta pupils in in grades IV and VII. The results of previous studies in this project have been reported by G. R. Conquest,[6] T. J. Reid[7] and Dr. H. T. Coutts and Dr. H. S. Baker.[8]

The study has been confined to an analysis of test results from a combined large urban and small urban sample. The grade IV sample contained 350 pupils (199 large urban and 151 small urban); while the grade VII contained 341 pupils (197 large urban and 144 small urban).

The tests used in the project were the California Language Test, Elementary, Form AA, for the grade IV sample and the California Language Test, Intermediate, Form AA, for the grade VII sample. These tests were developed jointly by Ernest W. Tiegs, Ph.D., and Willis W. Clark, Ed.D., and are published by the California Test Bureau—5916 Hollywood Boulevard, Los Angeles 28, California.

In determining the effectiveness of the California Language Test items as general achievement test items, two related factors were investigated: (a) the discriminating power of the test items, and (b) the difficulty of the items.

At present, a variety of item analysis procedures is being used in test construction. The statistics most frequently employed for the purpose are the critical ratio, the product-moment coefficient of correlation, and the biserial coefficient of correlation. The biserial correlation method is considered by many as being the most satis-

---

[6]G. R. Conquest, **A Survey of English Language Achievement in Grades Four and Seven in Selected Aberta Schools.**

[7]T. J. Reid, **A Survey of the Language Achievement of Alberta School Children in Relation to Bilingualism, Sex and Intelligence.**

[8]H. T. Coutts and H. S. Baker, "A Study of the Written Composition of a Representative Sample of Alberta Grade Four and Grade Seven Pupils," **Alberta Journal of Educational Research (June, 1955), pp. 5-18.**

factory measure of item validity or internal consistency. It measures the relationship "between each item and the total score, excluding the item in question".[9] This coefficient would probably find wider use if its computation were less labourious.

To approximate the biserial correlation coefficient, with less labour, a method was devised by T. L. Kelly[10] in which he demonstrated that the lowest and highest twenty-seven per cent of a sample are the most serviceable groups for use in item-analysis, even though the items are not all of fifty per cent difficulty.

In 1939, J. C. Flanagan,[11] using the method described by Kelley, devised a chart showing the values of product-moment coefficients corresponding to given proportions of successes in the upper and lower twenty-seven per cent of a sample. While the coefficients taken from Flanagan's chart are unaffected by item difficulty, they do not represent a linear function and therefore cannot be added, substracted or averaged.

To satisfy the requirement for an index of discrimination which is a linear function, Fredrick B. Davis prepared an item analysis chart containing discrimination indices derived from Flanagan's product-moment coefficients. Flanagan's coefficients, which range from —.93 to +.93, were transformed into Fisher's z's. Davis then multiplied the z values, which represent a linear function, by a constant to obtain item discrimination indices having a range from —100 to +100.

The Flanagan method and the Davis method, which require the same arrangement of data, were used for item analysis in this study. To determine the general pattern of item difficulty and average level of difficulty of sub-tests, the percentage of successes for each item was computed.

Employing these methods, the writer examined the elementary and intermediate tests for ineffective items and determined how well the tests are adapted to the language abilities of Alberta grade IV and grade VII pupils.

Item-Analysis Data

The item-analysis techniques described by Davis and Flanagan require computation of the proportion of successes on items in the twenty-seven per cent of the sample scoring the highest on the criterion and the twenty-seven per cent scoring the lowest.

9Davis, **op. cit.**, p. 9.

10T. L. Kelley, "The Selection of Upper and Lower Groups for the Validation of Test Items," **Journal of Educational Psychology** (Vol. XXX), pp. 17-24.

11J. C. Flanagan, "General Considerations in the Selection of Test Items and a Short Method of Estimating The Product-Moment Coefficient from Data at the Tails of the Distribution," **Journal of Educational Psychology** (Vol. XXX), pp. 674-680.

The proportion of successes is defined as the number of testees that know the answer to an item divided by the number of testees that respond to the item. The reader's attention is drawn to the fact that the word "know" is used in the definiion rather than the word "mark". This distinction is desirable whenever multiple-choice items are used because a certain number of testees, without knowing the answer, will mark the correct one by chance. To make allowance for this fact, wherever possible, a correction for chance has been employed.

The scoring procedure in Sections A and B of both tests reduces the factor of chance by allowing credit for an item when a testee may select any number of the possible responses, provided that the correct response is included in the selection. It is possible, therefore, for a testee to receive a perfect score by simply filling-in all response spaces. No penalty is imposed for over-capitalization or over-punctuation.

In Section B, of both tests, a place is provided for the testees to indicate that no punctuation is necessary, but such is not the case in Section A—Capitalization. Here, no capitalization is indicated by no response to an item. This results in an ambiguity in the interpretation of responses. No response may indicate that the testee refrained from marking an answer because of lack of knowledge or a lack of time. Or it might be interpreted to mean that the testee, after reading the item, decided that no capitalization was needed.

In the computation of the proportion of successes on items of Section A and B, of both tests, no corrections for chance were made. The proportions of successes in the remaining sections were corrected for chance.

Following the computation of the proportion of successes, item analysis charts were used to determine the discrimination index for each item.

The item-analysis chart, devised by Davis, is arranged in such a way that the percentage of succeseses of the twenty-seven per cent of pupils scoring highest on the test is located on the horizontal axis. The percentage of successes of the lowest twenty-seven per cent is located on the verical axis. By reading horizontally and vertically from the previously determined percentages. two indices may be located simultaneously, one immediately above the other. The upper index is the discrimination index of an item and the lower, a difficulty index.

The Flanagan chart is constructed in a similar manner, with the proportions of successes of the highest twenty-seven per cent placed on the horizontal axis and the proportion of successes of the lowest

twenty-seven per cent on the vertical axis. Product-moment coefficients of correlation may be located in the same manner as are the Davis indices of discrimination. The coefficients are given in multiples of five with each of the same value and sign being joined by a line. A coefficient, other than one of a multiple of five, must be estimated by its location between two adjacent lines.

Table I contains item-analysis data based on the test results from the California Language Test, Elementary. The data in Table II are based on the results from the Intermediate test.

In both Table I and Table II, R refers to the number of pupils responding to the items, and P the proportion responding to the item correctly. Item 5, 14 and 28 in Table I and 6, 12, 13, 20 and 24 in Table II require two responses each. The data for these items are numbered 1 and 2 in each table.

It will be noted that, for a few items, the proportion of successes is negative after a correction for chance had been made.

> A negative proportion may be interpreted to mean that fewer testees marked the correct answer than would have been expectede by chance, probably because many of them were misinformed rather than simply uninformed with respect to the point being tested.[12]

Where the signs of the two proportions used to enter the item-analysis chart are different, there is no direct method of arriving at an accurate index of discrimination. Davis states that, in his experience, serviceable indices can be obtained for items when the proportion of successes in the lowest twenty-seven per cent of the sample is negative by altering the proportion arbitrarily to the proportion that would result from one-half a testee having answered the item correctly.

In this study such a procedure was followed by entering the item-analysis chart at one per cent on the vertical scale when a negative proportion occurred. The discrimination index thus obtained is always an underestimate. This fact is noted in Tables I and II, by the addition of a plus superscript to the discrimination index, and Flanagan's "r".

Each item that correlates positively with the final score contributes to the final measurement, to some extent, regardless of how small the correlation might be. The higher the correlation the more efficient is the item for general achievement test purposes.

Items which exhibit either zero or negative correlations with the total score do not contribute to the final measurement and therefore have no value as general achievement test items.

---

[12]Davis, op. cit., p. 3.

TABLE I

ITEM-ANALYSIS DATA FROM THE CALIFORNIA LANGUAGE TEST (ELEMENTARY)

Item No.	Highest 27%		Lowest 27%		Davis Discrim. Index	Flanagan r
	R	P	R	P		
A. Capitalization						
1	92	.98	87	.93	21	.35
2	89	.95	65	.69	28	.43
3	81	.86	72	.77	8	.13
4	90	.96	56	.60	37	.54
5—1	93	.99	75	.81	35	.52
2	88	.94	67	.72	24	.37
6	73	.78	43	.46	21	.34
7	93	.99	74	.80	36	.55
8	90	.95	58	.64	34	.51
9	90	.95	58	.64	34	.51
10	92	.98	66	.73	35	.52
11	84	.89	67	.71	17	.27
12	71	.76	36	.41	23	.36
13	48	.51	20	.23	19	.31
14—1	82	.87	59	.67	17	.27
2	74	.79	31	.35	30	.45
15	85	.90	64	.73	17	.27
B. Punctuation						
16	83	.90	55	.59	24	.37
17	85	.89	56	.49	30	.45
18	72	.71	50	.41	19	.31
19	75	.79	39	.27	52	.70
20	54	.47	16	.17	22	.35
21	74	.74	39	.27	31	.46
22	87	.90	44	.47	33	.50
23	57	.51	19	.24	18	.29
24	89	.93	63	.62	29	.43
25	85	.89	44	.46	33	.50
26	84	.87	29	.18	52	.69
27	87	.89	54	.57	26	.40
28—1	61	.57	21	.08	39	.56
2	29	.30	5	.04	30	.46
29	81	.83	44	.45	27	.41
30	88	.89	46	.46	33	.50
C. Words and Sentences						
31	93	.98	78	.66	40	.59
32	94	1.00	85	.81	35	.52
33	93	.98	90	.91	17	.27
34	84	.79	84	.79	0	.00
35	94	1.00	91	.94	19	.31
36	93	.98	78	.66	40	.59
37	94	1.00	69	.47	57	.74
38	70	.49	42	—.11	55+	.71+
39	93	.98	89	.89	20	.32
40	93	.98	82	.74	34	.51
41	81	.72	67	.43	19	.32

TABLE I—*Continued*

Item No.	Highest 27%		Lowest 27%		Davis Discrim. Index	Flanagan r
	R	P	R	P		
42	90	.91	59	.26	49	.66
43	88	.87	53	.13	59	.72
44	85	.81	38	—.18	75+	.86+
45	87	.85	48	.04	67	.80
46	90	.91	69	.53	31	.46
47	93	.98	47	.04	76	.88
48	89	.89	72	.60	24	.38
49	85	.81	35	—.21	75+	.86+
50	64	.39	22	—.49	55+	.72+

The point at which a positively correlating item should be removed from a general achievement test cannot be stated with any finality. It depends upon the author's ability to produce another item which measures the same elements yet correlates more highly with the final score. As indicated in the study by Lawshe and Mayer,[13] reliability of measurement might be raised by removing seriously defective items or by adding items which exhibit high correlations with the total score.

Davis states:

> Items with discrimination indices above 20 will ordinarily be found to have sufficient discriminating power for use in most achievement and aptitude tests.[14]

Presumably, those with discrimination indices of less than twenty might be replaced as achievement test items. This constitutes the basis upon which items of the California Language Tests were judged in this study.

Based on the data summarized in Table I the following conclusions appear to be justified.

(1) Item 34 with a discrimination index of zero, should be removed from the test. This item fails to discriminate between levels of ability of testees in the groups of large urban and small urban Alberta grade IV pupils.

(2) Items 3, 11, 13, 14 (1), 15, 18, 23, 33, 35, 39 and 41, should be replaced, if possible. These items have discrimination indices of 20 or less.

(3) Items 37, 38, 43, 44, 47, 49 and 50 have the highest discrimination indices in the test. These items have indices of 55 or more.

[18] Lawshe and Mayer, op. cit., p. 277.

(4) Of a total of fifty-three items in the California Language Test, Elementary, twelve, or twenty-three per cent, fail to meet the requirements as suggested by Davis.

Table II contains the item-analysis data derived from the test results of the grade VII sample on the California Language Test, Intermediate.

TABLE II

ITEM-ANALYSIS DATA FROM THE CALIFORNIA LANGUAGE TEST (INTERMEDIATE)

Item No.	Highest 27%		Lowest 27%		Davis Discrim. Index	Flanagan r
	R	P	R	P		
A. Capitalization						
1	91	.99	84	.91	23	.37
2	92	1.00	72	.78	37	.57
3	84	.91	43	.47	35	.53
4	91	.99	87	.95	16	.25
5	92	1.00	77	.84	32	.48
6—1	88	.96	76	.78	24	.39
2	89	.97	76	.83	21	.34
7	88	.96	54	.59	37	.56
8	88	.96	77	.84	19	.31
9	88	.96	55	.60	37	.55
10	92	1.00	85	.92	22	.35
11	91	.99	69	.75	40	.58
12—1	92	1.00	73	.79	36	.54
2	92	1.00	76	.83	33	.50
13—1	83	.90	85	.92	— 3	—.05
2	89	.97	67	.82	24	.39
14	73	.79	54	.59	14	.23
15	67	.82	36	.39	29	.45
B. Punctuation						
16	71	.72	70	.74	— 1	—.02
17	91	.99	73	.79	36	.54
18	83	.90	58	.58	26	.41
19	70	.70	45	.40	19	.31
20—1	72	.73	46	.41	20	.32
2	66	.65	28	.16	34	.51
21	84	.90	62	.64	23	.37
22	87	.94	62	.64	29	.43
23	80	.79	47	.42	25	.39
24—1	75	.77	43	.38	27	.41
2	60	.57	22	.07	42	.60
25	87	.93	65	.72	22	.35
26	86	.92	59	.63	26	.41
27	89	.96	61	.66	32	.48
28	92	1.00	73	.86	30	.45
29	49	.42	24	.12	24	.38
30	90	.97	70	.86	19	.31
31	82	.89	59	.70	17	.27
32	64	.64	56	.64	0	.00
33	35	.26	11	—.06	41+	.59+

TABLE II—*Continued*

Item No.	Highest 27%		Lowest 27%		Davis Discrim. Index	Flanagan r
	R	P	R	P		
34	74	.85	52	.69	15	.23
35	81	.95	57	.77	21	.34
C. Words and Sentences						
36	92	1.00	83	.80	36	.54
37	86	.87	77	.67	17	.27
38	90	.96	81	.76	25	.40
39	89	.93	71	.54	34	.51
40	85	.85	57	.24	43	.61
41	91	.98	86	.87	22	.35
42	86	.87	72	.57	23	.37
43	69	.50	46	.00	55	.73
44	76	.65	57	.24	27	.41
45	92	1.00	86	.87	29	.45
46	91	.98	69	.50	50	.68
47	92	1.00	73	.59	50	.68
48	92	1.00	84	.90	25	.40
49	91	.98	75	.69	37	.55
50	92	1.00	77	.73	42	.60
51	92	1.00	77	.75	40	.58
52	91	.98	75	.70	37	.55
53	86	.87	51	.16	51	.69
54	88	.91	67	.52	32	.48
55	78	.70	33	—.25	68+	.81+
D. Parts of Speech						
56	87	.93	44	.35	46	.64
57	90	.97	55	.50	46	.64
58	81	.85	25	.09	57	.74
59	43	.33	16	—.03	46+	.64+
60	87	.93	60	.57	32	.48
61	34	.21	20	.02	30	.46
62	72	.73	12	—.09	69+	.81+
63	88	.95	34	.22	58	.75
64	82	.86	39	.29	40	.58
65	85	.90	36	.25	48	.66
66	57	.52	22	.05	41	.59
67	88	.95	28	.13	66	.80
68	89	.96	61	.59	37	.55
69	89	.96	41	.32	54	.72
70	92	1.00	59	.58	51	.69
71	8	—.14	7	—.15	0	.00
72	81	.86	29	.16	50	.68
73	79	.84	26	.12	52	.70
74	45	.37	20	.03	38	.56
75	89	.97	54	.49	47	.64

The following conclusions are based on the data presented in Table II and have reference to Alberta grade VII pupils of large urban and small urban areas.

(1) Items 13 (1) and 16 discriminate negatively and therefore are of no value as general achievement test items.

(2) Items 32 and 71, with discrimination indices of zero, fail to discriminate between grade seven Alberta pupils of high and low levels of ability and should be removed from the test.

(3) Items 4, 8, 14, 19, 20 (1), 30, 34 and 37 have discrimination indices of 20 or less. According to the suggestion of Davis, these should be replaced by items which have higher discriminating power.

(4) Items 43, 58, 62, 63 and 67 have the highest discrimination indices in the test. Their indices are 55 or more.

(5) Of the eighty items in the intermediate test, thirteen, or sixteen per cent, fail to meet the requirements of achievement test items as proposed by F. B. Davis.

Data on Item Difficulty

If the only discrimination indices used in selecting items for a test are based on the total score, it is desirable that items having the highest indices be chosen. But if maximum efficiency of measurement is to be attained the requirements of distribution and pattern of item difficulty must also be taken into account.

Items measuring ability at one level with high discriminating power, may have very low discriminating power at a different level. The fact is well known that the discriminating power of an item is greatest when it is used in a sample in which fifty per cent of the testees respond to the item correctly.

In the construction of achievement tests, where maximum discrimination among all testees is desired, the optimum shape of the distribution becomes a function of the intercorrelations of the items.

The fact that the average discrimination index for all items is low in both tests may be due to the fact that the items are measuring rather specific elements that are not highly intercorrelated. Therefore, it might be expected that the items would be distributed over a wide range of difficulty with the majority near the fifty per cent level of difficulty.

In this study, the percentage of successes is used as an index of item difficulty. It is the percentage of pupils in the sample who responded correctly to a test item. The percentage of pupils responding successfully, unsuccessfully and making no response are

summarized in Tables III and IV. The letter S refers to the percentage of pupils responding correctly, U refers to the percentage responding incorrectly and N refers to the percentage making no response.

TABLE III

PERCENTAGE OF SUCCESSES AND FAILURES ON ITEMS OF THE CALIFORNIA LANGUAGE TESTS, ELEMENTARY

Item No.	S	U	N
A. Capitalization			
1	94	6	0
2	86	14	0
3	69	31	0
4	88	12	0
5—1	89	11	0
2	86	14	0
6	74	26	0
7	93	7	0
8	74	26	0
9	78	22	0
10	88	12	0
11	87	13	0
12	63	34	3
13	45	51	4
14—1	72	24	4
2	50	46	4
15	77	19	4
Mean % of Success = 77			
B. Punctuation			
16	68	32	0
17	71	28	1
18	53	43	4
19	60	39	1
20	68	31	1
21	55	45	0
22	70	28	2
23	66	33	1
24	76	24	0
25	72	26	2
26	54	40	6
27	64	24	12
28—1	36	33	31
2	12	59	29
29	63	27	10
30	59	28	13
Mean % Success = 56			
C. Words and Sentences			
31	90	10	0
32	98	2	0
33	99	1	0
34	89	11	0
35	98	2	0
36	89	11	0
37	84	16	0
38	64	36	0
39	99	1	0
40	97	3	0
41	84	16	0
42	87	13	0
43	79	21	0
44	61	39	0
45	82	18	0
46	95	5	0
47	88	11	1
48	91	8	1
49	66	33	1
50	43	53	4
Mean % Success = 84			

Mean Percentage of Success = 74
(all items)

An examination of Table III reveals the following:

(1) The range of difficulty of the items is satisfactory but the test contains too many relatively easy items. With a mean percentage of success of 74, it contains too many items at the lower levels of difficulty and too few near the 50 per cent level.

(2) The punctuation test, with a mean percentage of successes of 56, provides the best distribution of item difficulty of the three sub-tests. However, it contains too few items at the lower levels of difficulty.

(3) Discriminations between levels of ability in the A and C sub-tests are likely to be less than optimal because of the low range of the item difficulty, and because of concentration of items at the lower levels of difficulty.

Table IV summarizes the percentages of pupils responding successfully, unsuccessfully and making no response to the eighty items of the California Language Test, Intermediate.

An examination of the levels of item difficulty in Table IV indicates that there is an adequate range but that the test, like the elementary, contains too many relatively easy items. The mean percentage of success for the test is 77.

The capitalization, and words and sentences sub-tests, each with a mean percentage of success of 86, have very similar distributions of item difficulty. Both sub-tests contain too many relatively easy items and have insufficient range of item difficulty for the sample of pupils used in this study.

The punctuation test, with a mean percentage of success of 78, contains too many items at the lower level of difficulty and too few near the 50 per cent difficulty.

Of the four sub-tests, the parts of speech test contains items covering the widest range of difficulty. There are, however, too many items at the lower levels.

Both the elementary and intermediate tests were used on samples of the lowest grades for which they were designed. The elementary test, intended for grades IV, V and VI, was administered to a grade IV sample, while the intermediate test, intended for grades VII, VIII and IX, was administered to a grade VII group.

Since the tests contain an unduly large number of relatively easy items for grade IV and grade VII pupils, it might be expected that the distributions of item difficulty would be still more seriously skewed if the tests were employed at their upper grade levels. This would result in less effective measurement of general achievement when the tests are administered to grades VI and IX pupils.

The current practice in test construction is to present items covering a wide range of difficulty, arranged in ascending order of difficulty from the very easy to the most difficult. This makes it possible for the less able pupil to respond to items within his level of ability without being discouraged early in the test by items of

TABLE IV

PERCENTAGE OF SUCCESSES AND FAILURES ON ITEMS OF THE CALIFORNIA LANGUAGE TEST, INTERMEDIATE

Item No.	S	U	N
A. Capitalization			
1	96	4	0
2	90	10	0
3	69	31	0
4	97	3	0
5	94	6	0
6—1	91	9	0
2	91	9	0
7	79	21	0
8	92	7	1
9	80	20	0
10	98	1	1
11	89	10	1
12—1	93	6	1
2	93	7	0
13—1	85	14	1
2	89	10	1
14	71	27	2
15	55	42	3
Mean % Success =86			
B. Punctuation			
16	74	26	0
17	94	6	0
18	85	15	0
19	64	35	1
20—1	72	28	0
2	48	50	2
21	81	18	1
22	79	19	2
23	72	26	2
24—1	69	29	2
2	46	52	2
25	86	11	3
26	82	15	3
27	86	11	3
28	91	5	4
29	37	59	4
30	90	5	5
31	81	12	7
32	60	31	9
33	20	69	11
34	69	16	15
35	79	6	15
Mean % Success = 78			

Item No.	S	U	N
C. Words and Sentences			
36	93	7	0
37	88	12	0
38	94	6	0
39	88	12	0
40	76	24	0
41	97	3	0
42	84	16	0
43	60	40	0
44	72	28	0
45	98	2	0
46	89	11	0
47	91	9	0
48	97	2	1
49	92	7	1
50	93	6	1
51	95	4	1
52	93	6	1
53	73	26	1
54	89	10	1
55	59	40	1
Mean% Success = 86			
D. Parts of Speech			
56	73	27	0
57	86	14	0
58	60	40	0
59	34	66	0
60	85	15	0
61	29	71	0
62	43	57	0
63	70	30	0
64	69	30	1
65	73	26	1
66	41	59	0
67	64	36	0
68	86	14	0
69	77	22	1
70	85	14	1
71	10	89	1
72	67	32	1
73	60	39	1
74	31	68	1
75	82	17	1
Mean % Success = 61			

Mean Percentage of Success = 77
(all items)

prohibitive difficulty. The most difficult items, at the end of the test, differentiate between pupils of superior ability.

Where an achievement test contains a number of sub-tests, an arrangement of items in ascending order of difficulty might be expected in each sub-test. They should begin with their easiest items and end with the most difficult.

An inspection of Table III reveals a tendency toward an increase in item difficulty in the sub-tests of the elementary test. This pattern of item difficulty is only slightly apparent from the test results of the sample employed in this study. The large proportion of items of approximately the same difficulty causes the pattern of item difficulty to be barely discernible.

A similar pattern of item difficulty in the intermediate test is somewhat apparent in only the capitalization, and words and sentences sub-tests. In the punctuation, and parts of speech tests, items are arranged in an irregular fashion with respect to item difficulty.

Items in the sub-tests of the elementary and intermediate tests may have been arranged in ascending order of difficulty for the normalizing sample but evidence of this arrangement for the sample in this study is minimal.

Conclusions

(1) Defects in the scoring procedure and in the method of recording answers in the capitalization and punctuation sections reduces the effectiveness of the test items. Equal credit is given to pupils marking only the correct response, and to others marking several, of which one is correct. There would appear to be a difference in the ability of these pupils but this is not always indicated by the final score.

(2) In both tests, the Davis discrimination indices are relatively low, indicating a lack of test homogeneity. As a result, some ambiguity in the interpretation of the total score may be expected.

The fact that items appear to be measuring specific elements which are not highly correlated may account for the lack of internal consistency in the test.

Because of their restricted content, the sub-tests may possess greater test homogeneity and yield relatively unambiguous scores. But before one places any credence in their results he should be assured of their reliability. The test authors of the California Language Tests have not provided this information.

(3) Perhaps the greatest factor limiting the effectiveness of the California Language Tests is the unduly large proportion of re-

latively easy items. The levels of item difficulty are not well adjusted to the abilities of Alberta grade IV and grade VII pupils in large urban and small urban centres. The large proportion of relatively easy items has the effect of reducing the power of the test to differentiate between pupils at the upper levels of ability.

Since the tests contain too many relatively easy items for the sample of grade IV and VII pupils, it might be expected that the distributions of item difficulty would be still more seriously skewed if the tests were administered to grades six and nine pupils.

(4) The California Language Tests are limited in their usefulness in that they measure but a few of the basic objectives of language instruction. Their content is restricted to the measurement of the use of capital letters and punctuation marks, and the recognition of complete sentences and parts of speech. No attempt is made to measure some of the broader outcomes such as the ability to organize and present material logically, to sustain an idea, to establish continuity of thought and to discriminate in the selection of words. For this reason they must be regarded as tests of language usage rather than comprehensive tests of language ability. This is a weakness which is common to the language portion of most of the current achievement tests.

BIBLIOGRAPHY

Anderson, John E. "The Effect of Item Analysis upon the Discriminative Power of an Examination," *Journal of Applied Psychology* 19 (1935), pp. 237-244.

Conquest, George R. *A Survey of Language Achievement in Grades IV and VII in Selected Alberta Schools.* Unpublished Master's Thesis, University of Alberta, 1954.

Coutts, H. T. and Baker, H. S. "A Study of the Written Composition of a Representative Sample of Alberta Grade IV and Grade VII Pupils." *Alberta Journal of Educational Research* 1:2 (June, 1955), pp. 5-18.

Davis, Frederick B. *Item-Analysis Data.* Cambridge, Mass., Harvard University, 1946.

Flanagan, John C. "General Considerations in the Selection of Test Items and a Short Method of Estimating the Product-Moment Coefficient from Data at the Tails of the Distribution." *Journal of Educational Psychology* XXX (1939), pp. 674-680.

Hawkes, H. E., Lindquist, E. F. and Mann, C. R. *The Construction and Use of Achievement Examinations.* Boston, Houghton Mifflin Company, 1936.

Kelley, T. L. "The Selection of Upper and Lower Groups for the Validation of Test Items." *Journal of Educational Psychology,* XXX (1939), pp. 17-24.

Lawshe, C. H. and Mayer, James S. "The Effect of Two Methods of Item Validation on Test Reliability." *Journal of Applied Psychology* 31 (1947), pp. 271-277.

Reid, T. J. *A Survey of the Language Achievement of Alberta School Children in Relation to Bilingualism, Sex and Intelligence.* Unpublished Master's Thesis, University of Alberta, 1954.